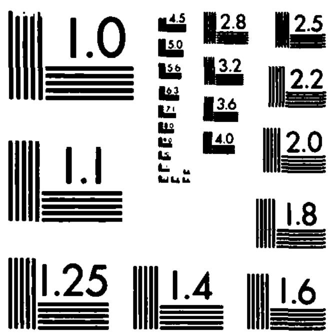
AD-A161 066 VALIDATION OF AIR FORCE HAZARD ASSESSMENT RATING 1/2
METHODOLOGY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON
AFB OH SCHOOL OF SYSTEMS AND LOGISTICS M C ANDERSON
UNCLASSIFIED SEP 85 AFIT/GEM/LSY/85S-1 F/G 6/20 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

VALIDATION OF AIR FORCE

HAZARD ASSESSMENT RATING METHODOLOGY

THESIS

Mr. Myron C. Anderson

AFIT/GEM/LSY/85S-1

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GEM/LSY/85S-1

VALIDATION OF AIR FORCE

HAZARD ASSESSMENT RATING METHODOLOGY

THESIS

Mr. Myron C. Anderson

AFIT/GEM/LSY/85S-1

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GEM/LSY/85S-1

VALIDATION OF AIR FORCE HAZARD ASSESSMENT RATING METHODOLOGY

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Engineering Management

Myron C. Anderson, B.T.

September 1985

Approved for public release; distribution unlimited

Preface

The purpose of this study was to validate and evaluate the Air Force Hazard Assessment Rating Methodology (HARM). The HARM model is used to evaluate past hazardous waste disposal sites. HARM predicts which hazardous waste sites warrant detailed investigation based on the potential risk to our health and environment.

The results of the Phase II investigation were used to validate the HARM model. At the time of this study the data from Phase II investigations were limited. Although the limited data restrict the direct application of this research, a method to improve the HARM model was developed and other valuable information was obtained.

The preparation of this thesis would not have been accomplished without assistance from others. I wish to thank Mr. Rich Murphy, my advisor, and Lt Col Al Tucker, my reader, for their assistance and direction. A special thanks should also go to Donna and Debbie for their endless typing and editing assistance.

To my wife Donna and sons Matthew and Steve, I express my deepest gratitude for their encouragement, patience, and understanding.

Myron C. Anderson

Table of Contents

List of Figures

List of Tables

Abstract

The Air Force uses the Hazard Assessment Rating Methodology (HARM) for the initial screening of uncontrolled hazardous waste sites in the Installation Restoration Program (IRP). This initial screening evaluates the potential health and environmental hazards associated with the site and determines if the site warrants further investigation. There is a definite need to properly evaluate these sites because investigative costs are high and it is important not to eliminate sites that need further investigation.

This research evaluated results from actual Phase II investigations. Discriminant analysis was used to improve the HARM model's ability to properly evaluate sites.

The results indicate that current HARM procedures correctly predict which sites need further investigation and which do not only 68% of the time. Through the application of classification techniques developed in this study, the preditive capability was increased by over twenty percentage points so that approximately 90% of the sites were correctly classified. Before the refined model is applied more Phase II data are needed for final evaluation and testing of the new model.

VALIDATION OF AIR FORCE
HAZARD ASSESSMENT RATING METHODOLOGY

I. Introduction

Overview

In 1981, the Environmental Protection Agency (EPA) reported that there were 14,098 hazardous waste generators who generated 264 million metric tons of regulated hazardous waste (8:5). Prior to 1976 when the Resource Conservation and Recovery Act (RCRA) was passed, there was little control or regulation of hazardous waste disposal. Many materials in fact were not recognized as hazardous and therefore were disposed of in what was then thought to be a "satisfactory" manner. Today many of those disposal sites are causing health and environmental problems. It is estimated that over 20,000 sites in the United States contain hazardous substances and that as many as 2,000 of these sites pose a significant hazard to public health and the surrounding environment (9:757).

The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) of 1980 was passed in part to identify and clean up those hazardous sites generated in the past. CERCLA, which is commonly known as the "Superfund" bill, states that federal agencies are responsible for the identification and clean up of their own uncontrolled hazardous waste sites. The United States Air Force (USAF) is currently in the process of doing this under the Department of Defense (DOD) directed Installation Restoration Program (IRP).

The process of identifying uncontrolled hazardous waste sites receives considerable public attention, often raising emotions and controversy. Additionally, the USAF effort to identify and clean up these sites is expensive. The DOD cost of the entire Installation Restoration Program is expected to cost between $5 billion and $10 billion and take more than 10 years to complete (12:9-10). The cost of the USAF IRP through 1991 alone is expected to exceed one billion dollars (7). With this amount of public attention and investigative costs in the millions of dollars, it is imperative that sites are properly identified and hazards accurately assessed. Costs increase significantly when sites progress beyond the initial identification and assessment phase and into the investigative and clean up phases. Consequently, USAF sites improperly identified as hazardous sites when, in fact, they are not can result in considerable unnecessary costs to the Air Force. But of equal or greater concern is the possibility of not properly identifying all sites that are hazardous. This situation could result in direct adverse impact on public health or the environment as well as be a discredit to the USAF.

To properly identify uncontrolled hazardous waste sites and to assess site risks or hazards, the USAF conducts an initial records search at a base. Site hazards are evaluated using the USAF Hazard Assessment Rating Methodology (HARM). HARM prioritizes sites based on their estimated hazards to public health, welfare, and the environment. Sites with a high HARM score require more extensive investigation while sites with a low HARM score are eliminated from further consideration.

Background

EPA Superfund Program. On 11 December 1980, the President signed into law the Comprehensive Environmental Response, Compensation, and Liability Act of 1980. The act established a program to deal with the problem of uncontrolled hazardous waste sites. Sites are considered uncontrolled and hazardous if they are releasing or threatening to release hazardous substances into the environment under unregulated conditions. CERCLA also established procedures for identifying and reporting these sites as well as new hazardous substance spills or releases. The financial responsibility and liability of owners, operators, and users (both generators and transporters) of hazardous waste sites are explained in some detail.

Hazardous waste sites vary considerably in size and characteristics. They include landfills, chemical disposal pits, drum and tank storage or disposal areas, contaminated areas as a result of spills or explosions, and surface impoundments, virtually any place where hazardous materials were deposited. The hazard associated with a site can vary significantly depending primarily on the amount of hazardous material and the degree of hazard associated with that material, as well as where the site is located and how well the hazardous material is contained.

A substance can be considered dangerous or hazardous if exposure could result in: acute toxicity, carcinogenic, mutagenic, or teratogenic effects; devaluation or loss of property; or the destruction of endangered flora or fauna or their habitat (6). Also dangerous are substances that are ignitable, reactive, or explosive. Some examples of these types of substances include organic solvents, pesticides, heavy metals, strong acids and bases, salts, munitions, fuels, and many other organic substances. The substances create a hazard when they contaminate the soil, surface or ground water, if they give off toxic or explosive gases, or if they are in an unconfined area where people or animals can come in direct contact with them.

In addition to identifying and reporting uncontrolled hazardous waste sites, Section 104 of CERCLA requires remedial action where necessary (6:0703). The remedial action may entail field investigation, containment, removal, treatment, and recovery or proper disposal of the hazardous substances and contaminated resources. Relocation of nearby residents or other restrictions may also be required. If necessary, an alternate water supply may need to be developed. Site rehabilitation and monitoring are also normal requirements. CERCLA established a trust fund to help pay for these activities at sites where responsible parties cannot be identified or are unable to pay. Where federal agencies are responsible they are required to fund their own remedial actions. To identify the worst sites that require remedial action and possible federal funding, EPA developed a Hazard Ranking System (HRS). The system is used to evaluate uncontrolled hazardous waste sites that are candidates for the National Priority List (NPL). The NPL now contains over 500 sites for remedial action and funding from the special trust fund (11).

DOD Installation Restoration Program. Because CERCLA gave responsibility to federal agencies for their own remedial action activities, including funding, DOD established its own clean up program. Actually the DOD program was established in 1975 when the Army was in the process of getting rid of excess property. The IRP was established to identify and evaluate "contaminated lands and facilities which were or might become excess to DOD's needs" (13:1).

On 24 June 1980, in response to RCRA requirements and in anticipation of CERCLA, DOD established a comprehensive program for all DOD installations by issuing a Defense Environmental Quality Program Policy Memorandum (DEQPPM 80-6), which was later updated by DEQPPM 81-5 issued 11 December 1981. The DOD Installation Restoration Program was established:

> to (1) identify and evaluate suspected problems associated with past hazardous material disposal sites located on DOD installations and (2) control the migration of hazardous environmental contamination from these sites [13:1].

In order to clarify federal agencies' responsibilities under CERCLA legislation, the President issued Executive Order 12316 dated 14 August 1981. Responsibilities delegated to the Secretary of Defense for DOD facilities and vessels were as follows:

1. Response actions for the removal of contaminants,

2. Remedial actions to solve the problems caused be hazardous waste pollution,

3. Investigations including monitoring, surveying, and testing for hazardous waste and hazardous waste sites,

4. Planning, legal, fiscal, economic, engineering, architectural, and any other studies or investigations as necessary for proper response actions (13:12).

<u>Air Force Installation Restoration Program</u>. The USAF provided its initial policy guidance based on DOD guidance in December of 1980 and actually started its program early

in 1981. The program received technical guidance and direction as well as oversight from the Air Force Engineering and Services Center (AFESC). The USAF, because of its primary defense mission, has long been using and handling a wide variety of toxic and hazardous materials similar to any other large industrial complex.

The Air Force IRP is divided into four phases. They are briefly described as follows:

1. Phase I - Installation Assessment (Records Search). Phase I is designed to identify and assess those past disposal sites that may pose a hazard to public health or the environment as a result of uncontrolled hazardous waste contamination at the site. To evaluate a site, all available relevant information is collected from records, interviews, and on site observations. Only existing information is used as no analytical sampling of the suspected hazardous waste or the surrounding environment is conducted during this phase. To evaluate all the information collected during Phase I, the USAF developed the Hazard Assessment Rating Methodology (HARM). The HARM score provides a basis for determining which sites warrant further investigation (Phase II) and which sites require no further action because they are considered safe and pose no significant threat to public health, welfare, or the environment. Sites investigated that clearly pose no significant potential hazard are not even rated by the HARM

model. Phase I becomes a basic background document for the Phase II study when the study is necessary.

2. Phase II - Confirmation/Quantification. Phase II is to confirm, define, and quantify by preliminary and comprehensive investigative surveys, the presence (or absence) of contamination and the extent of contamination. This phase identifies sites where remedial action will be required in Phase IV. If a site requires immediate remedial action, such as removal of abandoned drums, the action can proceed as soon as it is identified. Research requirements identified during this phase will be addressed in Phase III of the program.

3. Phase III - Technical Base Development. The purpose of this phase is to develop a sound data base upon which to prepare a comprehensive remedial action plan. Phase III includes implementation of research requirements for new and improved technology to identify, investigate, and mitigate or clean up sites. A Phase III requirement can be identified during any other phase of the program.

4. Phase IV- Operations/Remedial Actions. This phase includes the preparation and implementation of the remedial action plans that will mitigate, control, or eliminate the hazardous waste site. It is the final phase and often referred to as the clean up phase.

When the USAF started the Installation Restoration Program, EPA had not yet published its official Hazard

Ranking System (HRS) which it uses to evaluate sites. The USAF, in its desire to start identifying and cleaning up hazardous waste sites and to show environmental leadership, started evaluating its sites. To do this it was necessary to have some means of evaluating the information collected. The USAF therefore developed an evaluation method which was based on a rating system developed by JRB Associates, Incorporated (1). The JRB system was one of the intial systems proposed to EPA for its use. The USAF modified the JRB system slightly so that it would provide meaningful ratings for various types of Air Force sites including small landfills, fire training areas, and spill sites. This initial USAF evaluation and rating system was later refined to what is now called HARM, as explained in Chapter III. This revision was made by people who were experienced in evaluating health and environmental hazards and had been using the initial system for approximately 6 months at USAF installations. The HARM model is currently used for all USAF sites evaluated during Phase I.

The HARM model is a system which evaluates potential hazardous waste sites and provides a relative ranking for each site depending on its degree of hazard. This allows the USAF to compare sites at all its installations. Sites that require further investigation under Phase II can be identified and prioritized based on the degree of hazard. The site ranking or priority is derived from a numerical scoring system. Site characteristics such as; hazardous waste types and quantities, pathways available for contaminant migration,and proximity to areas that would be adversely impacted by contaminant migration are considered. Each of the characteristics are weighted and combined in an algorithm to produce an overall site rating score. Since evaluation of HARM is the main thrust of this thesis, it will be discussed in more detail in Chapter III.

Problem Statement

The USAF HARM system has received a fair amount of attention for several reasons. The public is concerned about hazardous materials, and the news media frequently discusses problems associated with hazardous substances. Places like Times Beach and Love Canal are familiar names because of the news coverage of the environmental and health problems at these uncontrolled hazardous waste disposal sites. The USAF does not want this type of attention and is trying to identify and clean up its own problem sites. States and local communities are applying pressure, even insisting that the USAF conduct sampling and monitoring at all of their sites no matter how small the potential hazard. This would be a costly endeavor and in some cases an outright waste of financial resources. On the other hand, the USAF is also concerned about health, safety, and the environment. It must minimize the possibility of failure to investigate sites that need investigation. The decision to investigate or not is made primarily on the basis of a HARM rating obtained during the IRP Phase I records search. It is clearly the critical input in deciding whether or not the site needs further investigation.

Second, the HARM system is not as familiar as HRS, which the EPA uses to evaluate its sites. Although HARM is a refinement of an earlier USAF system, it has not been revalidated or formally evaluated since it was put into use in 1982. Because HARM plays such a critical role in the decision process, there is a definite need to substantiate that it is providing the correct information for the correct decisions, thus accomplishing its purpose.

The HARM rating is used to help determine which sites warrant further investigation based on the potential for hazardous waste contamination and adverse impacts. Validation of the HARM model will show how effective the current model is in identifying sites that have a high probability of hazardous waste contamination. Phase II investigative results showing the sites where significant hazardous waste contamination was found can be used to validate the model.

Without this validation study there is no assurance that all of the sites that need investigation are being properly identified. Validation of the HARM model will provide support for its use, increasing confidence in HARM both within and outside the USAF. It will help insure that the USAF is identifying sites in an economical and environmentally safe manner. In addition, statistical analysis may result in refinement of the model which would result in better identification of sites needing investigation or a simplified model requiring less information but with equal predictive capability. This would lead to a better understanding of which variables are critical in predicting the sites with uncontrolled hazardous waste contamination.

Research Objectives

General. The main thrust of this research was to validate and refine the effectiveness of the current HARM system. Validation and refinement of HARM will help insure proper identification of uncontrolled hazardous waste sites that need further investigation.

Specific. The specific objectives of this research are listed as follows:

1. Documentation of the Phase I HARM ranking system's effectiveness in identifying sites that need further investigation and sites that require no further action.

2. Improve HARM's ability to discriminate between sites that need further investigation and those that require no further consideration.

3. Simplify the HARM model without sacrificing

effectiveness. The goal was to have equal or improved predictive capability with less information.

4. Identify any types of sites or situations where HARM is less effective and attempt to isolate their unique characteristics, if any.

Scope. Currently available IRP Phase I and Phase II reports were used to provide the necessary data for analysis. No attempt was made to collect additional site data or information. The statistical analysis was limited by the amount of Phase II investigative data available. Some Phase I information has already been compiled for sites on over 70 bases, but less than one-half that amount is available from Phase II reports and none of it has been compiled or categorized.

Although the literature review covered various methodologies and theories on risk assessment, there was no extensive comparison of the HARM model with other existing models. Models thought to be similar often have a different purpose or require different information and are therefore difficult to compare.

Research Questions

To solve the stated problems and meet the research objectives the following investigative questions were used to guide the research:

1. What does IRP Phase II data show about HARM's

ability to identify hazardous waste sites that need investigation?

2. Can HARM's predictive capability be improved?

3. Can HARM maintain its effectiveness in predicting contaminated sites with less information?

4. Does HARM work equally well for all types of sites and situations?

II. Literature Review

There is more and more awareness of the risks associated with everyday life. This awareness has increased as regulatory agencies have tried to evaluate and control the risks of numerous hazards. Risk assessment and management are an integral part of controlling environmental health hazards. Because of the increased use of risk assessment, there is an increased need for scientists, engineers, policymakers, and the general public to have a better understanding of risk assessment and risk management. This literature review will look at information relating to environmental health risk assessments.

The literature review will address the problem of defining and quantifying risk and its relationships to other functions. A significant part of the review will deal with defining what risk assessment is and what problems are associated with its use. Risk assessment approaches will be discussed conceptually, along with relationships to risk management, regulatory agencies, and the public. The review is intended to provide a better understanding of risk assessment in the environmental health field and not to be an exhaustive review of all literature relating to risk.

Discussion of Terms and Concepts

Risk. Risk is a term that is familiar to most people. However, that fact in itself can be a problem since different people think of risk in different ways. Before any discussion of a topic can occur it must be properly defined. The terms "risk" and "hazard" are synonyms in our everyday language, but in a scientific context they are defined differently. Hohenemser, from the Department of Physics, Clark University, made the following distinction:

> Hazards are threats to humans and what they value, whereas risks are quantitative measures of hazard consequences that can be expressed as conditional probabilities of experiencing harm. Thus, we think of automobile usage as a hazard but say that the lifetime risk of dying in an auto accident is 2 to 3 percent of all ways of dying [16:379].

A different conceptual definition of risk is provided by Kahan, of the Department of Psychology, University of Southern California, who relates risk to the decision making process. A decision is required when a choice must be made between several alternatives, each of which has consequences associated with it. If it is known what the consequences are and that they will occur, then the decision is made under certainty; and if the consequences are unknown, the decision is made under uncertainty. Decisions are made under risk when the consequences can only be estimated or when the consequences occur with a known probability. Consequences can be either good or bad and are often referred to as benefits or hazards (costs) respectively. Thus, hazards describe negative consequences of a decision and risk describes the situation where the consequences or hazards are estimated or when their probability of occurrence is known (19:1-2). Salem, of the Rand Corporation, highlights the problem of defining risk by summarizing three distinctly different views of risk:

> 1. In a decision-analytic sense, risk refers to the fact that the consequences of choosing an alternative are not known with certainty, but instead can be expressed as probabilistic outcomes. In this sense, no reference needs to be made about the positive or negative effect of the consequences.
>
> 2. By contrast, the popular view of risk emphasizes the probability of a potential harm and focuses on the probability of that harm without regard to the (negative) magnitude. The benefits in this sense are completely ignored.
>
> 3. Between these two viewpoints lies a third definition of risk--as the product of the probability and magnitude of each undesirable possible outcome, integrated or summed over all undesirable outcomes [28:5].

Risk Assessment. The general process of evaluating risk is often referred to as risk assessment. The term "probabilistic risk assessment" (PRA) is used frequently by the nuclear power industry and Nuclear Regulatory Commission (NRC). Levine, formerly of the NRC, describes PRA as "a method of predicting the likelihood and consequences of reactor accidents" (21:14). William D. Ruckelshaus, Environmental Protection Agency Administrator, defines risk assessment as follows:

> Risk assessment is the use of a base of scientific research to define the probability of some harm coming to an individual or a population as a result of exposure to a substance or situation [28:5].

As shown in this definition, the primary emphasis is on the risk to human health, but it does not preclude risks to our environment. The definition also stresses probability more than magnitude. Risk assessment uses data from experiments and other scientific data bases to estimate probable human health effects resulting from exposure to hazardous substances.

Risk Management. Risk is a controversial issue for numerous reasons. One reason is that in the past there has been no distinction between risk assessment and risk management. The problems and complexity of dealing with risk in relation to human health exposure has been addressed by the National Academy of Sciences (NAS). They recommend that risk assessment and risk management be separate functions within regulatory agencies to the extent possible. Miller, an editor for Environmental Science and Technology magazine, reported that the NAS distinction between risk assessment and risk management is as follows:

> Risk assessment is only one aspect of risk management - the scientific component. Whereas risk management involves choices between the broader social and economic policy issues, risk assessment considers only scientific problems and does not involve socioeconomic matters. The scientific findings and judgments embodied in risk assessment should be explicitly distinguished from the political, economic, and technical

> considerations that influence the design and choice of regulatory strategies. By separating the two, there will be for the first time a clear demarcation between science and the policy aspects of regulatory decision making [24:200A].

To further define these different functions Miller provides this additional clarification:

> Risk assessment uses factual information from laboratory and other sources to estimate possible human health effects from exposure to hazardous substances. It uses a scientific base, for example, dose-response relationships to define the health effects of exposure of individual populations to hazardous materials or situations.
>
> Risk management, on the other hand, weighs policy alternatives and selects the most appropriate regulatory action; risk management integrates the results of risk assessment with engineering data and with social, economic, and political concerns to reach a regulatory decision [24:200A].

Risk management decisions are thought by some people to be similar to cost/benefit management decisions (14:12).

The merit in separating the scientific function from the controversial regulatory function is easy to see. Morton Corn, director of the Johns Hopkins University's Department of Environmental Health Sciences and former head of the Occupational Safety and Health Administration, as reported in Chemical and Engineering News, made the following observation:

> Separating risk assessment from risk management won't decrease the controversy of regulation. But having the risk assessment separated will make clear why we are doing what we are doing, and increase public understanding [3:4].

After the NAS and Mr. Ruckelshaus came out in support of separating risk assessment and risk management, different opinions started appearing. Hanson reported an opinion that "risk assessment cannot be purely scientific, that it must be recognized as being at least partially subjective" (14:12). Subjectivity in any analysis opens it up to controversy which would defeat the purpose of separating risk assessment and risk management. Mr. Ruckelshaus stated that it is difficult to keep risk management factors out of the risk assessment process (27:35).

Site Contamination. Since site contamination is a major part of this research, the term needs clarification. For the purpose of this research, a site is any location where hazardous waste has been deposited or was suspected of being deposited. It includes the entire disposal area even though the hazardous waste may only be in part of the site. Contamination occurs when the hazardous waste migrates beyond the site where it was deposited. Either surface water or ground water normally provides the pathway for contaminant migration. A contaminated site or site contamination means that hazardous waste is migrating beyond the immediate site boundaries.

Regulatory Problems

Approach. A recognized regulatory problem is the lack of a uniform approach to risk assessment. Mr. Ruckelshaus, as reported in Science magazine, stated that current regulations designed to control public health risks are not consistent across government agencies and that "a common statutory framework for dealing with environmental risks is needed" (29:1026).

An NAS recommendation was explained by Miller as follows:

> It advised the federal government to adopt a package of standard procedures for estimating health effects from hazardous chemicals, drugs, and food additives. It also called for uniform risk assessment guidelines... [24:200A].

Other organizations such as the International Agency for Research on Cancer and the Office of Technology Assessment also have proposed uniform risk assessment guidelines (25:62). A uniform approach would give all the regulatory agencies and scientists some common ground or a frame of reference for further discussion. It would eliminate some of the controversy resulting from different agencies trying to establish different regulatory standards for the same substance (24:200A).

There has also been controversy over the way risks are regulated within the same agency. The EPA does not apply the same standards for risks from new products and industries as it does for old products and industries. The risks of new products are screened very carefully. Many old products are routinely used and accepted even though the risks associated with them are known, and they would not pass today's risk screening standards. This in effect is a double standard for risk between old and new products. A uniform approach to risk assessment for comparing the risks of various alternatives would eliminate this problem (17).

Methodology. Currently there are many approaches and methods to assess risk. Dr. John W. Hernandez, Deputy Administrator of the EPA, according to the Journal of Environmental Health, described the risk assessment process as answering two questions:

> 1. How likely is the adverse effect to occur?
>
> 2. What is the magnitude of the public health impact at ambient exposure levels, if the event does occur [15:63-64]?

He further states that the first step is to determine from the data whether or not a substance is hazardous, i.e., can it be classified as a suspected carcinogen. The evidence or data supporting the determination can be strong (based on human data) or weak (based on short-term, high dose animal studies). He also states that the second step is the risk quantification stage and that EPA tends to over-estimate the risk because of the high level of uncertainty and concern for public health.

One method of risk assessment that has received considerable use by the NRC and nuclear industry is probabilistic risk assessment (PRA) (21). Levine describes the PRA process as follows:

The heart of PRA is contained in logic models known as event trees and fault trees. Event trees describe initiating events such as pipe breaks, and subsequent successes or failures of the systems designed to cope with them. They contain accident sequences that could cause the nuclear fuel to be severely damaged, and the subsequent processes that could damage the containment building and release radioactivity to the environment. Fault trees describe the ways in which the systems involved in the accident sequence could fail, and provide estimates of the frequencies of such failures [21:41].

Although the PRA is widely used, it is not without its critics. MacKenzie quotes several Advisory Committee on Reactor Safeguards memos as follows:

The large uncertainties inherent in PRA are well recognized and acknowledged in the proposed policy statement. These uncertainties make the use of PRA in decision making (which occurs already within the NRC) subject to large differences in the results obtained by different groups of analysts for the same accident scenarios. These uncertainties also permit abuse of the methodology to obtain a result which supports a predetermined position by selective choice of data and assumptions. The claims for PRA concerning its ability to assess public-safety risks are little more than a sham that will hide the fact that the basis for safety will always depend upon the judgement of a few individuals [22:36-37].

The lack of sufficient scientific data and high levels of uncertainty are significant risk assessment problems (24:20CA). The NRC recognized the effect of uncertainty, and according to Solomon, of the Rand Corporation, there is no standard way to compensate for the level of uncertainty in regulatory decision making (30:1-3). He went on to make the following recommendation:

> Implementation of the proposed numerical guidelines for reactor safety must take into account this fundamental limitation of PRA and its associated uncertainty. In probabilistic risk assessments made in conjunction with safety goals, the underlying assumptions should be disclosed....[30:3].

In a recent speech, Mr. Ruckelshaus expressed concern about the effect of limited data and high uncertainty. He also indicated several times that there is often considerable pressure from the public's desire for certainty. Scientists and regulatory agencies must continue to resist the temptation to imply certainty where none exists (27).

HRS and HARM Comparison

As previously mentioned, models are difficult to compare and the HRS and HARM models are also slightly different. The HRS model is designed to evaluate only the worst sites for possible federal funding and requires some initial sampling and analysis. The HARM model requires no initial sampling, evaluates any site with potential hazards, and is used to determine if sites warrant further investigation. But even with these differences they are both used to evaluate uncontrolled hazardous waste sites.

The EPA felt that the USAF and other services should use the HRS model, but the USAF felt the HARM model was tailored for their purposes and best served their needs. To help clarify the situation Engineering-Science, Inc., at the request of the USAF, compared site ratings using both models at sites on four USAF bases (5:1-4). The site scores derived from HRS and HARM were significantly different for the same sites. In general, the HARM scores were higher and had a wider range than the HRS scores. This can be attributed to the different algorithms used in each method. Some of the criteria considered in each model were different and criteria weightings were not always the same. Even with these differences there was one important comparison which showed a strong similarity. The rank order correlation between sites rated under each model was high. With only a few exceptions sites were ranked in the same priority sequence under both methods. Both models have their advantages and disadvantages, but the HARM model provides a broad range of site scores and allows for a wide variety of types of USAF sites.

Summary and Application

The literature shows that there are many ways to evaluate or assess risk. There is no consensus on which methodology should be used. Each method appears to be tailored for its specific purpose. The approach of separating risk assessment from risk management appears to have merit but the application of this principle is difficult.

HARM does not attempt to separate risk assessment from risk management. It uses factual site information and then assigns a weight or value judgement about socioeconomic matters. Most of the methodologies of risk assessment reviewed considered the probability of an event occurring and the magnitude of the impact in one way or another. The HARM system considers the magnitude of hazards to populations and environments and determines the most likely pathway that could allow hazardous contaminant migration to the populations or environment.

III. USAF Hazard Assessment Rating Methodology (HARM)

Overview

CERCLA requires EPA to evaluate waste sites and to develop a national priority list (NPL) of sites for remedial actions. Section 105 of CERCLA (6) spells out the minimum criteria to be considered such as, the population at risk, the specific hazards associated with the substances present, the potential for contamination of drinking water supplies, direct human contact, destruction of sensitive ecosystems, and other appropriate factors. The USAF HARM evaluation system as well as EPA's system must include these criteria. HARM uses only available information with no preliminary analytical sampling to determine if Phase II investigation (sampling) is needed and to prioritize, based on potential hazard the follow-on investigation. Phase II confirms if there is contamination at the site or not. The USAF does not use HARM to rate a site unless it determines there is sufficient hazardous waste present to have a potential for contamination. Figure 1 provides a general decision tree of the IRP Phase I process and shows how sites are deleted from the HARM model rating process, based on initial Phase I information.

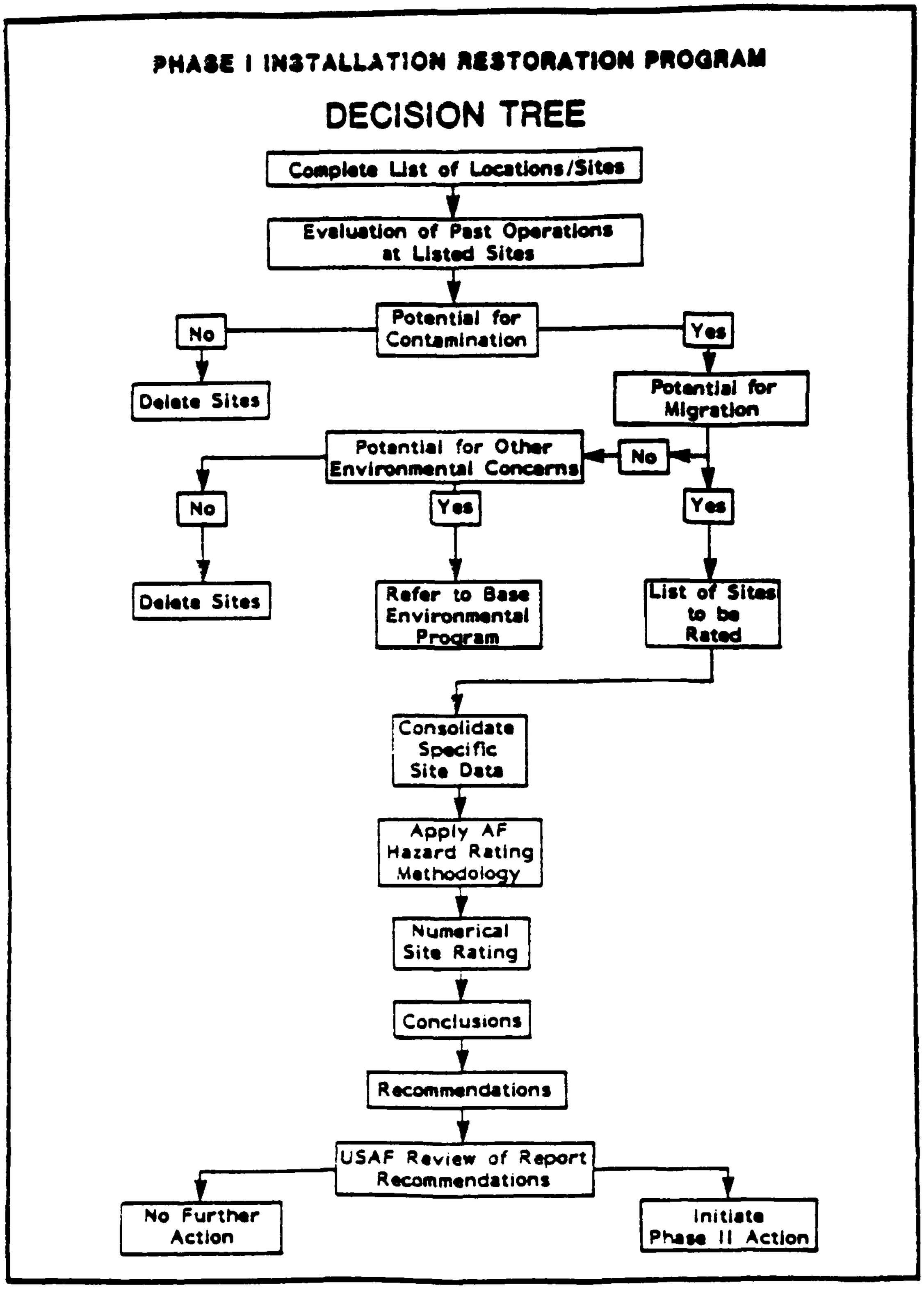


Fig. 1. Decision Tree (18)

HARM Model Description

General. The model uses a numerical scoring system to rank sites. The input data for the model is readily available from actual site visits, existing records, and interviews with potential contributors to and operators of the site. Specific scoring criteria is given in an effort to make required subjective judgements more objective for consistent evaluations.

HARM considers four categories of site characteristics: 1) potential receptors of the contamination, 2) waste material characteristics, 3) potential pathways for contaminant migration, and 4) any management efforts to contain or control the hazard. In essence HARM considers the "who, what and how" for each disposal site. A flow chart is provided in Fig. 2 to show category relationships and general scoring methodology. The receptors, waste characteristics, and pathways scores are averaged and multiplied by the waste management score to arrive at the overall HARM score for that site. Figure 3 provides an example of the rating form and its use.

An entire description of HARM methodology and the scoring criteria is contained in Appendix A. To aid in understanding the HARM rating process an example will be used to explain the four major categories.

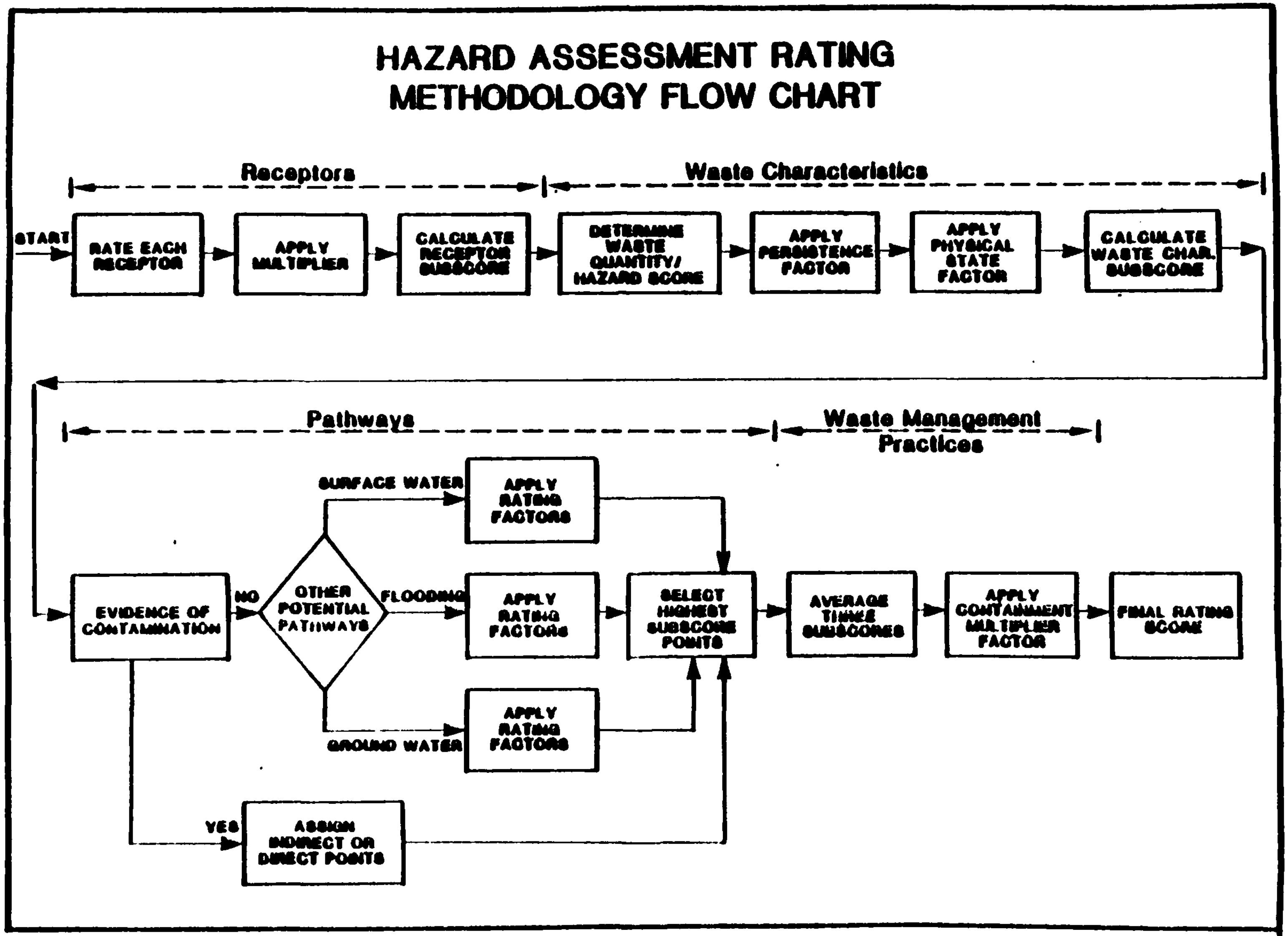


Fig. 2 Flow Chart (18)

HAZARDOUS ASSESSMENT RATING FORM

NAME OF SITE: No. 10--JP-5 Fuel Spill No. 2

LOCATION: AF

DATE OF OPERATION OR OCCURRENCE: 1981

OWNER/OPERATOR: AF

COMMENTS/DESCRIPTION: 21,000-gallon JP-5 fuel spill

SITE RATED BY: K. Cable

I. RECEPTORS

	Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
A.	Population within 1,000 feet of site	2	4	8	12
B.	Distance to nearest well	1	10	10	30
C.	Land use/zoning within 1 mile radius	3	3	9	9
D.	Distance to reservation boundary	2	6	12	18
E.	Critical environments within 1 mile radius of site	1	10	10	30
F.	Water quality of nearest surface-water body	3	6	18	18
G.	Ground-water use of uppermost aquifer	0	9	0	27
H.	Population served by surface-water supply within 3 miles downstream of site	3	6	18	18
I.	Population served by ground-water supply within 3 miles of site	0	6	0	18
			Subtotals	85	180

Receptors subscore (100 x factor score subtotal/maximum subtotal) 47.2

II. WASTE CHARACTERISTICS

A. Select the factor score based on the estimated quantity, the degree of hazard, and the confidence level of the information.

1. Waste quantity (S = small, M = medium, L = large) L
2. Confidence level (C = confirmed, S = suspected) C
3. Hazard rating (H = high, M = medium, L = low) M

Factor Subscore A (from 20 to 100 based on factor score matrix) 80

B. Apply persistence factor
Factor Subscore A x Persistence Factor = Subscore B

80 x 0.8 = 64

C. Apply physical state multiplier

Subscore B x Physical State Multiplier = Waste Characteristics Subscore

64 x 1.0 = 64

Fig. 3. Rating Form (18)

III. PATHWAYS

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
A. If there is evidence of migration of hazardous contaminants, assign maximum factor subscore of 100 points for direct evidence or 80 points for indirect evidence. If direct evidence exists then proceed to C. If no evidence or indirect evidence exists, proceed to B.				
			Subscore	80
B. Rate the migration potential for three potential pathways: surface-water migration, flooding, and ground-water migration. Select the highest rating, and proceed to C.				
1. Surface-water migration				
Distance to nearest surface water	3	8	24	24
Net precipitation	2	6	12	18
Surface erosion	2	8	16	24
Surface permeability	2	6	12	18
Rainfall intensity	2	8	16	24
		Subtotals	80	108
Subscore (100 x factor score subtotal/maximum score subtotal)				74.1
2. Flooding	0	1	0	3
	Subscore (100 x factor score/3)			0
3. Ground-water migration				
Depth to ground water	2	8	16	24
Net precipitation	2	6	12	18
Soil permeability	1	8	8	24
Subsurface flows	0	8	0	24
Direct access to ground water	1	8	8	24
		Subtotals	44	114
Subscore (100 x factor score subtotal/maximum score subtotal)				38.6

C. Highest pathway subscore

Enter the highest subscore value from A, B-1, B-2, or B-3 above.

Pathways Subscore **80**

IV. WASTE MANAGEMENT PRACTICES

A. Average the three subscores for receptors, waste characteristics, and pathways.

Receptors	47.2
Waste Characteristics	64.0
Pathways	80.0
Total 191.2 divided by 3 =	63.7
	Gross Total Score

B. Apply factor for waste containment from waste management practices

Gross Total Score x Waste Management Practices Factor = Final Score

63.7 x 1.0 **64**

Fig. 3. Continued (18)

Receptors Category. The receptors category considers potential impact on public health, welfare, and the environment in the vicinity of the site. It considers who and what can be affected by contaminant migration and exposure. There are nine receptor factors, each of which is rated in one of four levels (0-3). The criteria for each level is specified so that anyone with the same information will rate the receptor factor the same. The zero level rating means the receptor is at minimal risk. A three level rating means the receptor is at the maximum risk. Each of the nine receptor factors has a weighting multiplier assigned to it. The receptor factors with the highest weightings are the distance to the nearest well, the critical environment within one mile, and the ground water use of the uppermost aquifer. The rating factor level is multiplied by the weighting to get the receptor factor score. All of the receptor factor scores are added, totaled and the percentage of the maximum possible score is calculated. This is the final receptor category score. The receptors category helps to define the magnitude of potential hazard. In the example (Fig. 3) the receptor factor subtotal score is 85 out of a possible 180 for a final receptors category score of 47.2.

Waste Characteristics Category. The waste characteristics category score is computed in a different way. The initial category score is based on three factors:

the waste quantity, the degree of hazard associated with the waste, and the confidence level of the waste characteristic information. The hazardous waste quantity has three levels--small, moderate, and large--with specified amounts for each category. The degree of hazard rating also has three levels--high, medium, and low--with specified criteria for each. The degree of hazard category is further broken down into three subfactors: toxicity, ignitability, and radioactivity. The highest of the three subfactor ratings is used to indicate the degree of hazard category. The subfactor ratings are not recorded on the form, only the degree of hazard category. The confidence factor only has two levels, confirmed and suspected, which are based on the reliability and confidence of the information. Again, the criteria for the levels are defined to allow for consistent evaluations.

The three waste characteristics factors and their associated levels can result in 18 different combinations. Each combination has an assigned value which becomes the initial category score when that combination occurs. The highest initial category rating of 100 is for a confirmed large quantity of waste with a high degree of hazard, while the lowest initial category score of 20 is for a suspected small quantity of low hazardous waste. The initial category score is then adjusted twice, once for persistence of the waste substance and once for its physical state (solid, semisolid, or liquid). The final waste characteristic score is the initial category score multiplied by the persistence factor and then the physical state factor.

In the example the initial category score of 80 is based on the combination of a confirmed (C), large (L) quantity of medium (M) hazard waste. The waste characteristics matrix in Appendix A assigns a value of 80 for this combination. The initial score is then multiplied by a persistence factor of 0.8 and a physical state factor of 1.0 for a final waste characteristics category score of 64. The waste characteristics category score aids in defining the magnitude of potential hazard and the probability of contamination.

<u>Pathways Category</u>. The pathways category assesses the potential pathways for contaminant migration and assigns a pathways category score based on the most probable pathway. The most probable pathway is assumed to be the one with the highest pathway subscore. The pathway subscores are computed using the same approach used in the receptor category calculations. The approach considers three pathways (surface water, flooding, and ground water) creating three pathway subscores. Two other subscores based on evidence of contamination are considered along with the pathway subscores to obtain the highest pathway category score. If there is direct laboratory analysis indicating contamination. a subscore of 100 is assigned and this

becomes the pathway category score because it is the highest possible score. If there is indirect evidence of contamination, such as a visible leachate but no actual analysis, then a subscore of 80 is assigned. This becomes the pathway category score if it is higher than the other three pathway subscores. Only the highest score is used and there is no adjustment for more than one high pathway subscore.

The example shows a surface water pathway subscore of 74.1, a flooding subscore of 0, and a ground water subscore of 38.6. The surface water pathway shows the highest rating and would be used as the final category pathway category score except that there is indirect evidence of contamination which gives a higher score of 80. Therefore, the pathway category score is 80. The pathway and waste characteristic scores primarily define the probability and magnitude of contamination.

Waste Management Practices Category. This fourth and final category is not really a category but only a final management factor. The three category scores are averaged and then multiplied by a waste management factor to arrive at the final HARM score for the site. This approach considers management controls along with containment and monitoring provisions and adjusts the initial HARM score to reflect these management practices. There are only three management factor levels (1.0, 0.95, 0.10) used, limiting the options for this powerful adjustment to the total HARM score.

The site in the example has no management practices. Therefore, the management factor is 1.0 and the initial HARM score of 64 becomes the final HARM score. If a site is managed with total containment and monitoring the initial HARM score is reduced by 90%. This in effect eliminates the site from further consideration because the hazard is being controlled. But if there is only limited management control, containment, and monitoring then the HARM score is reduced by only 5%. This adjustment compensates for some site management even though the site is not completely controlled. The higher the final HARM score the greater the need for follow on investigation, Phase II.

Summary

The HARM algorithm is primarily a linear equation with several adjustments for first order interaction. Interaction takes into account the combined effects which are multiplicative and not just additive. The total HARM score is multiplied by the management factor to compensate for the overall interaction effect on the need for further investigation. The other place where interaction is considered is in the waste characteristics category. The interaction between the waste quantity, the degree of hazard of a waste and the confidence in that information is considered by assigning a waste characteristics score based on their combined effect. This interaction is not strictly multiplicative and is introduced by the structural matrix feature. The waste characteristics score is again adjusted for interaction because the physical state and persistence affect the potential hazard associated with the waste.

IV. Methodology

Overview

Model validation, evaluation, and refinement can be accomplished in several ways. The initial USAF hazardous ranking system was already revised once, as explained in Chapter I. The process relied on expert judgement to adjust the initial model so that the final HARM score more closely reflected the independent judgement. Without direct objective measurements this approach is a common internal validation procedure (10:128-129). In fact, EPA tested its HRS model by using an independent panel of experts to rank a group of sites and then compared the results with the HRS scores (4).

Now that more objective Phase II information exists, it can be used to validate HARM. The general objective was to see how effective the HARM model was in predicting which sites were contaminated and needed further investigation and to see if this predictive capability could be improved.

HARM estimates the risk or degree of hazard associated with an uncontrolled hazardous waste site. As explained in Chapter III, relevant information is combined into an overall HARM score which is used to determine if further investigation is warranted or not. The higher the score the greater the need for further investigation.

Data

A major consideration in any research project is the data that is used. This is especially true when the research involves statistical analysis such as that used in this research project. The source of the data was from published Phase I Records Search reports and Phase II Confirmation reports. In a few cases, additional information or clarification was obtained from interviews with people responsible for the reports, or from USAF IRP complete site status reports. All the bases included in the research are listed in Appendix B along with the complete data set. In all, data from 108 sites on 14 USAF bases were used. The total population of USAF sites projected to require Phase II investigation is expected to exceed 1000 sites on over 150 installations.

The research data set represented only about ten percent of this total population primarily because more Phase II data was not available. The research sample was also not representative of the types of Air Force bases being studied, with 11 of the 14 bases in either the Tactical Air Command or the Air Force Logistics Command. It is difficult to estimate the effects of the data limitations. Any model developed should be checked against sites within other Air Force Commands before it is accepted for general USAF application. This precaution is necessary because different Air Force Commands have different missions, handle different hazardous material, and have different management procedures. Therefore, their site characteristics may not be adequately considered by a HARM model based primarily on data from just two USAF Commands.

Another word of caution about the data is that it uses ordinal scale measurements. The assignment of a numeric rating level only indicates its rank order and no more. For example, a rating level of two for the depth to ground water factor does not indicate that it is twice as close as a rating level of one. The use of ordinal scale data for this research is generally not considered a problem. In fact, "a great many statistics assume no more than an ordinal level of measurement" (26:5). The dependent validation variable used in this research was the site classification. The independent or predictor variables are the site characteristics.

Predictor Variables. The HARM site characteristics were originally selected because they were felt to be good predictor variables. They include all the site characteristics used in the HARM model plus several additional waste site characteristics. The predictor variables from the HARM model and their associated weights are shown on the rating forms (Figure 3). Table I lists the additional predictor variables that will be considered. The additional predictor variables were selected to see if they could improve the HARM model's predictive capability.

Table I

Additional Predictor Variables

Predictor Variable
Type of Site
Landfill
Fire Protection Training Area
Spill Site
Waste Disposal Site (includes structures used for hazardous wastes, surface impoundments and burial sites)
Drum Storage
Radioactive Waste Area
Leaking Underground Storage Tanks or Pipes
Oil and Water Separators
Waste Treatment Plants
Munition Disposal Site
Others (includes pesticide areas, discharge areas and drainage systems)
General Hazardous Waste Class
Volatile Organics
Petroleum, Oils and Lubricants
Heavy Metals
Radiological Substances
Polychlorinated Biphenyls
Sludge
Munitions
Others (includes Pesticides, Herbicides, Fungicides, unknown types of Hazardous Chemicals, Refuse - general and industrial)
General Disposal
Above Ground
Below Ground
Other or Both

Site information that was not available or relevant in predicting the need for further investigation was not considered.

Site characteristic factor levels were also considered reliable. With the given criteria, any investigator could consistently come up with similar rating levels for each site characteristic. There could be some minor differences because subjective judgement is involved in some of the variables. Also, objective information about a site characteristic that is near a criteria break point between factor level scales can easily be rated in the wrong level because of a small error in the objective information.

Classification Variable. To validate the HARM model requires objective information concerning what has been predicted or estimated. This objective information must itself be a good indicator of the model's validity. To be a valid indicator the information must be relevant, reliable, and free of bias (10:130). Each of these attributes is discussed in this section.

HARM in part predicts the probability of contamination at a site. The Phase II site investigation provides direct relevant validating information on whether or not the site is contaminated, which is the primary input to the Phase II site classification. In general a site is considered contaminated if there is hazardous waste contamination beyond the immediate site boundary. All the information from the Phase II site investigation is used to classify the site as requiring no further action, long term monitoring, or remedial action. There is no detailed criteria for each classification except that a site requiring remedial action must have a contaminant concentration in excess of a health or environmental standard or guideline (2). The no action classification is assumed to mean that no contamination was found or that the contaminant was at a sufficiently low concentration, causing no significant hazard or adverse impact.

The long term monitoring (LTM) classification may reflect several situations. If the investigator is not confident or is unsure of the results of the investigation, the person may recommend long term monitoring. This situation would probably occur most frequently when no significant contamination was found but a safe conservative approach would be to recommend long term monitoring. This is not considered a good reason for this classification. The investigator should request more sampling and then properly classify the site. The logical and preferred reason for recommending the long term monitoring is that even though the investigation shows no significant contamination and the investigator is confident of that information, the person still feels the probability for contamination is high and/or the potential adverse impacts are significant if contamination does occur. The relevance of the three classifications to the HARM score must be examined and clarified.

The Phase II results indicate if contamination is occurring at the site and if so, at what concentrations. If significant contamination is found the site must be classified as requiring remedial action. If, on the other hand, no significant contamination is found, the site must be classified as requiring no action or long term monitoring. Only the portion of the HARM score that predicts the probability of contamination can be directly related to whether or not contamination was found. This would be a good relationship but the Phase II reports usually do not provide this information directly. Instead they report the site classification. The relationship between the final HARM score which reflects the need for further investigation and the Phase II site classification may not be as direct as desired but it is considered relevant for HARM validation. The problem is that a bias in the validation and refinement will occur if the site classification does not reflect everything HARM is measuring.

If the validation variable only reflects contamination, then only contamination variables will be important in predicting the need for further investigation. But HARM was designed to measure the need for further investigation by estimating the probability of contamination and the potential adverse impacts to receptors. Using only contamination as the validation variable is only part of the validation relationship and would result in built in bias against the importance of receptor factors.

Although site classification recommendations are based primarily on Phase II technical findings (contamination), they are also influenced by political and regulatory pressure as well as the proximity of the site to receptors. Because of this it was felt that the site classification data did not have significant amount of bias and that the Phase II classification could be used to validate the Phase I HARM model recommendations.

HARM classifies sites into two groups, those requiring further investigation and those that can be eliminated from further consideration. When the Phase I final HARM score does not clearly indicate to which group the site belongs, independent judgement is used to make that decision. That is why some sites with scores in the 40's were recommended for further investigation and some in the low 50's were not recommended for further investigation.

The relationship between the two Phase I HARM recommendations (investigate, no action) and the three Phase II classifications (required action, long term monitoring, no action) was not entirely clear, especially with the limited Phase II classification criteria. There was some question as to which Phase I group the long term monitoring was related. It was decided that long term monitoring was a form of required action or investigation and therefore, both it and the required action classification were directly related to the Phase I investigation recommended group. For analysis purposes these two Phase II classifications were combined into one group called required action. The Phase I and Phase II no action groups were clearly related.

The reliability of the site classification information is also an important factor in the validation process. The Phase II classifications were felt to be fairly accurate and consistent but there are potential problem areas. The entire Phase II process takes about four years to complete and because of this there are only a few bases with final Phase II reports (2). Even though there are only a few final Phase II reports, there are many interim reports. An initial Phase II interim report reports the results of an initial investigation designed to confirm site contamination, if it exists. Subsequent Phase II investigations are primarily concerned with characterizing, quantifying, and defining the extent of contamination as necessary. They also complete the confirmation of contamination where necessary.

Because only a few final Phase II reports were completed, site classification information was primarily obtained from the Phase II initial confirmation reports. Most of the sites were classified based on these initial reports but some sites needed further investigation. Only sites that could be classified with the initial confirmation reports were used, although subsequent investigations sometimes provided more site classification information. The site classification rarely changes in the final report so the interim reports were considered very reliable.

The other potential reliability problem area dealt with interpretation of the initial Phase II confirmation report results. Some reports did not clearly indicate the classification in which the site should be placed. Even though the researcher tried to be objective, there was an inclination to favor the no action classification for marginal sites. Therefore, some subjective bias may have crept into the classification process.

Phase II investigations generally follow the Phase I recommendations unless conditions change or there is regulatory pressure to investigate more sites. This means only a few sites in the group recommended for no further investigation actually have Phase II data. Of the 108 sites studied only 17 were from this group.

The sample distribution of sites among Phase II classification groups was satisfactory. There were 32 sites in the no action group, and 76 sites in the required remedial action group. The distribution among groups for the entire population of sites rated under HARM would probably have a higher percentage of sites in the no further investigation or no action group. Even though the sample distribution among groups may have been slightly different than that for the population, the sample was considered to have adequate representation for the no action and the required action groups. The statistical criterion for discriminate analysis which was used in this research is to have more sites in the smallest group than there are site characteristic predictor variables in the discriminant function (31:299).

Analysis Technique

To meet the research objectives several analysis techniques were followed. Descriptive statistics and discriminant analysis were the primary approaches used for data analysis. To perform the statistical research analysis, subprograms in the Statistical Package for Social Sciences (SPSS) release 9.1 were used (26).

Frequency analysis was the approach used on the ordinal scale data. The purpose was to gain insight into the HARM model variables. The mean, mode and range for variables in each of the groups were determined and differences were reviewed.

Discriminant Analysis. Since discriminant analysis was the primary statistical tool used for the research, a discussion of its features and their application is appropriate. Discriminant analysis is a statistical method for developing discriminant functions which can be used to assign observations to different groups. A data set which contains observations whose group membership is known is required to develop the discriminant functions. Each discriminant function is a linear combination of a set of discriminating or predictor variables. The predictor variables selected should measure characteristics that differ among groups. The coefficients assigned to the predictor variables are computed to maximize the difference in scores between groups. The maximum number of discriminating functions that can be derived is the lesser of one less than the number of groups or the actual number of discriminating variables (23:7-3). This research has a maximum of two groups and many discriminating variables. Therefore, the maximum number of discriminating functions is one.

The power of the discriminant function as a classifier of new cases (sites) can be evaluated by constructing classification functions from the discriminant function and then classifying cases with known group membership. To classify cases the value of the classification function for each group is calculated and the new case is assigned to the group with the highest value. The rate of correct classification provides an index of discriminating power for the set of variables used.

Classification functions are developed from Mahalanobis distance measurements. Mahalanobis distance is a generalized measurement of the distance from an individual case to a group centroid as measured in standard deviations from that centroid (20:44). A group centroid is the spatial representation of the group's mean for each of the predictor variables (20:16). The case is assigned to the group with the closest (the smallest Mahalanobis distance) group centroid. During the development of the discriminant function and when adding new discriminating variables, the objective is to maximize the distance between the closest groups.

The canonical correlation also aids in determining the importance of the discriminant function by telling how closely the function and the group variables are related (31:297). The higher the canonical correlation coefficient the better the association between the discriminant function and the groups. A value of zero means there is no association, and a value of one represents the maximum correlation. A large coefficient would indicate that the discriminant functions are able to accurately discriminate between or among groups. Therefore, the canonical correlation coefficient is a measure of the "power" of the discriminate functions (20:36-37). For the purpose of this research it is desirable to have a canonical correlation of at least 0.70 to show that there is a good association between the discriminant functions and the groups.

Another statistic used to evaluate the results of the discriminant analysis was the chi-square test of significance. The hypothesis being tested is that there is no difference between or among the groups under investigation. The possibility of finding significant results where none should exist is a consequence of relying on sample data. Random sampling procedures may produce data sets which are not good representations of the population. A P-value is the probability of obtaining the results of our discriminant analysis given the hypothesis is true. Therefore, a low P-value would indicate that the hypothesis is not true, which is the desired result. To help assure confidence in the discriminant functions developed the minimum P-value criterion was set at 0.01

Also important is the determination of the relative importance of each variable in the discriminant function. To do this, variable coefficients must be properly standarized for comparison. The larger the coefficient the greater the discriminating power of the variable.

Application

The first research objective of evaluating the HARM model's effectiveness in properly classifying sites was accomplished by comparing existing information. The approach was to compare the Phase I HARM predicted site grouping with the actual site classification based on the Phase II investigations. The percent of properly classified sites in each group and the overall percent correctly classified was computed. This was the measurement technique used to evaluate the effectiveness of the HARM model in predicting which sites needed further investigation.

The second objective was to improve HARM's ability to discriminate between sites that needed further investigation and those that required no further investigation. Discriminant analysis using the site characteristics as predictor variables and Phase II classifications as the validating classification variable, was the analytical tool used to meet this objective. The approach was to start with minor simple changes in the HARM model and then work toward more complex changes. Finally, additional site information was included.

The first step was to use the final HARM score as the only predictor variable. This is different than just comparing the Phase I HARM predictions because a single point or score was determined as the break point between groups with no judgement involved.

The next step was to use only the HARM category scores as predictor variables. After that analysis, the procedure was changed by ignoring all special features of the HARM algorithm and using all the raw information used in the HARM model. The information consisted of 28 factor scores representing site characteristics. Then to see if the special features in the HARM algorithm added any discriminating ability to the discriminant function, each of the HARM category scores was tested by using them in place of the factor variables they represented.

Based on the classification results at this point, special features that were beneficial would be simplified or improved if possible. Then new or different structural features and interaction between variables would be tested. Once the best classification function was obtained by using only the original HARM site information, additional available information was included in an attempt to further improve the new model's classification ability.

The third objective was to simplify the model without sacrificing its ability to classify sites. The approach was to select the best independent predictor variables based on their discriminating power. The program used a stepwise procedure to evaluate each variable according to its ability to discriminate between groups (26:235). The MAHAL stepwise evaluation method was used. This procedure seeks to maximize the Mahalanobis' distance between the closest two groups. By using the stepwise procedure "a reduced set of variables will be found which is almost as good as, and sometimes better than the full set" (26:447). Only the variables with the most unique discriminating power entered the model. If a variable duplicated the discriminating information of a variable already in the function, the new variable was not included. The result of this procedure was an optimal variable set being selected.

The last objective was to evaluate the sites that were misclassified in an effort to identify any unique characteristics. The approach was to see how far the individual site characteristic predictor variable values deviate from the associated mean and frequency values for the appropriate group. This will provide insight into possible model shortcomings or unique characteristics that may not be encountered at most hazardous waste sites.

Once the best possible model was obtained, it was tested for accuracy by using split-sample validation. Because the classification function is based on the sites it is trying to classify, there is a built in bias toward increased effectiveness in correctly classifying sites. To compensate for this, about twenty percent of the sites were excluded from the model development and then used to evaluate the classification function. There is a potential problem with this approach. Since using a data subset further reduces the sample size, it may affect the reliability of the classification function. That is why about 80% of the sites were used to develop the classification functions and only 20% were used to test them. The split-sample was stratified to maintain group proportions.

V. Results and Analysis

Descriptive Statistics

Descriptive statistics provide insight into the data used in this research. Site classification profiles were developed using frequency analysis to obtain the mode and range for each variable. Both the mode and the second most frequent value along with their associated percentages are provided. Mean values are used for interval level category scores and the final HARM score. Since the research approach combined the long term monitoring and the remedial action classes into one group, a profile of this new action required group was necessary. The other group profile was for the no action group. Also, a composite profile of all the sites in this research was included for completeness. The three profiles are presented in Tables II through IV. The information provided in these profiles is generally self-explanatory, but there are several significant results which need highlighting.

Total Data Evaluation. Although little difference was observed in the variable ranges of each group, it is interesting to note that the composite data had three site characteristics which did not use their entire range. The minimum coded values for fainfall intensity, depth to

Table II

Profile of All Sites

Site Variable	Mode/%	2nd/%	Code Range	Actual Range
Population	0/43%	1/34%	0-3	0-3
Distance to Well	3/52	1/19	0-3	0-3
Land Use	2/65	0/18	0-3	0-3
Distance within Base	3/51	2/33	0-3	0-3
Critical Environment	0/47	2/31	0-3	0-3
Surface Water Quality	1/45	2/23	0-3	0-3
Ground Water Use	0/42	1/27	0-3	0-3
Population Use of SW	0/88	3/7	0-3	0-3
Population Use of GW	3/69	1/16	0-3	0-3
Waste Quantity	1/40	2/34	1-3	1-3
Confidence Level	2/51	1/49	1-2	1-2
Waste Hazard Rating	3/86	2/12	1-3	1-3
Persistence	1/71	.9/15	.4-1	.4-1
Physical State	1/82	.75/15	.5-1	.5-1
Direct Contamination	0/94	1/6	0-1	0-1
Indirect Contam.	0/81	1/19	0-1	0-1
Distance to SW	2/44	0/20	0-3	0-3
Surface Erosion	0/55	1/31	0-3	0-3
Surface Permeability	1/47	0/46	0-3	0-3
Rainfall Intensity	2/49	3/34	0-3	1-3
Flooding Zone	0/71	1/23	0-3	0-3
Depth to Ground Water	3/59	2/28	0-3	1-3
Net Precipitation	2/43	0/20	0-3	0-3
Soil Permeability	3/48	2/35	0-3	1-3
Subsurface Flows	0/35	1/29	0-3	0-3
Direct Access to GW	0/68	3/20	0-3	0-3
Waste Management	1/76	.95/24	.95-1	.95-1

	Mean	Standard Deviation		
Receptors Score	48.3	15.2	0-100	25-91
Hazardous Waste Score	57.1	21.8	4-100	4-100
Pathways Score	71.7	16.8	0-100	28-100
Final HARM Score	58.2	10.7	0-100	29-85

GW-Ground Water
SW-Surface Water

Table III

Profile of Phase II No Action Sites

Site Variable	Mode/%	2nd/%	Code Range	Actual Range
Population	1/53%	0/38%	0-3	0-3
Distance to Well	3/38	0/28	0-3	0-3
Land Use	2/63	3/22	0-3	0-3
Distance within Base	3/38	2/34	0-3	0-3
Critical Environment	2/50	0/44	0-3	0-3
Surface Water Quality	1/50	2/31	0-3	0-3
Ground Water Use	0/63	3/22	0-3	0-3
Population Use of SW	0/97	3/3	0-3	0-3
Population Use of GW	3/66	0/25	0-3	0-3
Waste Quantity	1/66	3/19	1-3	1-3
Confidence Level	1/59	2/41	1-2	1-2
Waste Hazard Rating	3/88	2/9	1-3	1-3
Persistence	1/66	.8/19	.4-1	.4-1
Physical State	1/91	.5/6	.5-1	.5-1
Direct Contamination	0/100	1/0	0-1	0-0
Indirect Contam.	0/94	1/6	0-1	0-1
Distance to SW	3/53	1/31	0-3	1-3
Surface Erosion	0/56	1/31	0-3	0-2
Surface Permeability	1/63	0/34	0-3	0-2
Rainfall Intensity	3/41	1/31	0-3	1-3
Flooding Zone	0/75	1/25	0-3	0-1
Depth to Ground Water	3/63	1/19	0-3	1-3
Net Precipitation	3/41	2/38	0-3	0-3
Soil Permeability	2/38	3/34	0-3	1-3
Subsurface Flows	2/38	0/31	0-3	0-3
Direct Access to GW	0/79	3/17	0-3	0-3
Waste Management	1/75	.95/25	.95-1	.95-1
	Mean	Standard Deviation		
Receptors Score	42.3	13.4	0-100	26-87
Hazardous Waste Score	52.3	24.9	4-100	4-100
Pathways Score	63.7	15.0	0-100	28-100
Final HARM Score	52.1	9.0	0-100	29-71

GW-Ground Water
SW-Surface Water

Table IV

Profile of Phase II Required Action Sites

Site Variable	Mode/%	2nd/%	Code Range	Actual Range
Population	0/45%	1/26%	0-3	0-3
Distance to Well	3/58	1/17	0-3	0-3
Land Use	2/65	0/20	0-3	0-3
Distance within Base	3/57	2/33	0-3	1-3
Critical Environment	0/49	2/24	0-3	0-3
Surface Water Quality	1/43	0/21	0-3	0-3
Ground Water Use	1/36	0/33	0-3	0-3
Population Use of SW	0/84	3/9	0-3	0-3
Population Use of GW	3/70	1/18	0-3	0-3
Waste Quantity	2/42	3/29	1-3	1-3
Confidence Level	2/55	1/45	1-2	1-2
Waste Hazard Rating	3/86	2/13	1-3	1-3
Persistence	1/56	.9/16	.4-1	.8-1
Physical State	1/79	.75/20	.5-1	.5-1
Direct Contamination	0/91	1/9	0-1	0-1
Indirect Contam.	0/75	1/25	0-1	0-1
Distance to SW	3/63	2/21	0-3	0-3
Surface Erosion	0/54	1/30	0-3	0-3
Surface Permeability	0/51	1/41	0-3	0-3
Rainfall Intensity	2/58	3/32	0-3	1-3
Flooding Zone	0/70	1/22	0-3	0-3
Depth to Ground Water	3/58	2/32	0-3	1-3
Net Precipitation	2/45	1/26	0-3	0-3
Soil Permeability	3/54	2/34	0-3	1-3
Subsurface Flows	0/37	1/36	0-3	0-3
Direct Access to GW	0/51	3/30	0-3	0-3
Waste Management	1/76	.95/24	.95-1	.95-1
	Mean	Standard Deviation		
Receptors Score	50.9	15.2	0-100	25-91
Hazardous Waste Score	59.1	20.3	4-100	23-100
Pathways Score	75.1	16.4	0-100	42-100
Final HARM Score	60.8	10.3	0-100	43-85

GW-Ground Water
SW-Surface Water

ground water, and soil permeability were coded one instead of zero, which would indicate that the low end of the range for these characteristics was not used.

The direct access to ground water characteristic had 23 missing values due primarily to information being unknown when the Phase I study was conducted. The missing values were properly coded as zero, which would indicate no evidence of direct access risk. This allowed those sites with missing values for the direct access to ground water variable to remain in the analysis.

Phase II Group Profile Differences. Although the significance of group differences was evaluated by the discriminant analysis technique, several general observations are helpful at this point. The HARM model provided specific criteria for assigning ratings for each variable. However, the criteria were based on subjective estimates of relative importance with respect to potential site hazards without being validated against actual site conditions.

Consequently, when the model was applied in the field, some of the variables received ratings that tended to cluster at the high end of the scale without utilizing the full range of possible values. Likewise, other variables were almost always assigned values at the low end of the scale. This does not mean that these variables cannot be used to differentiate between groups. As long as the ratings associated with the required action group are generally higher than the ratings associated with the no action group, these variables can still provide important information about group differences irrespective of the limited range of assigned ratings. It is the differences between ratings for the two groups which become important for classification purposes and not the magnitude of the rating itself.

To evaluate group differences, the difference between mean values for site characteristic variables in the no action and the required action groups were compared. A variable was considered to have a positive group relationship if the mean value was at least 0.4 higher in the required action group compared to the no action group. Mean values that had less than the 0.4 difference between groups were considered neutral. If the variable had a 0.4 mean difference in the opposite direction (with the no action group higher) then the relationship was considered negative.

A negative variable contradicts the logic employed in establishing the HARM rating criteria. The ratings associated with the no action group are generally higher than the rating associated with the required action group. Even though the negative variables provide a basis for differentiating between the two groups, the apparent contradictions still need to be explained.

The mean difference criteria was based on a rating scale of zero to three. Therefore, the criteria for evaluating differences between means for variables which employed a different rating scale had to be modified accordingly. Table V lists the mean values for those variables which met the difference criteria. Group profiles, presented in Tables III and IV, provide information on the ratings assigned to all of the HARM variables.

Based on the difference between mean responses, six site characteristics had significantly higher means for the required action group than for the no action group. This difference in mean response indicates that the behavior of these variables supports the HARM relationship theory. Consequently, it is probable that these variables should appear as significant discriminators between the two groups.

Positive variable relationships were identified in each of the three major categories (receptors, waste characteristics, and pathways). In the receptors category, the distance to the nearest well and the distance within the base boundary showed a positive group relationship. The single positive relationship representing the waste characteristics category was the waste quantity variable.

The pathways category had three variables with positive relationships. The mean values for direct and indirect evidence of contamination do not meet the positive relationship criteria because this type of evidence is usually not available. However, they were considered to have a positive variable relationship because of group frequency differences. None of the seven sites with direct evidence of contamination were in the no action group, as would be expected. Only two of the twenty-one sites with indirect evidence of contamination were in the no action group. The direct access to ground water characteristic is somewhat subjective but it also showed a positive variable relationship.

Table V

Mean Profile Comparisons

Site Variable	Phase II Groups No Action	Phase II Groups Action	Phase I Groups No Action	Phase I Groups Action
Distance to Well	1.56	2.20	1.77	2.06
Distance within Base	2.06	2.46	1.71	2.46
Waste Quantity	1.53	2.00	1.47	1.93
Direct Contamination	0.00	0.09	0.00	0.08
Indirect Contam.	0.06	0.25	0.06	0.22
Direct Access GW	0.58	1.25	0.18	0.90
Land Use	1.94	1.57	1.53	1.70
Net Precipitation	2.00	1.36	1.71	1.55
Subsurface Flows	1.44	1.07	0.53	1.32
Soil Permeability	2.06	2.44	2.59	2.26

GW-Ground Water
SW-Surface Water

It is interesting to note that about half the site characteristics had a neutral variable to group relationship. The waste hazard rating characteristic and the waste management factor show no difference in their mean values and little frequency difference between groups. The hazard rating is high for all sites, which indicates that most of the sites in this research contained wastes considered to have a high degree of associated hazard. Management at hazardous waste sites showed no difference between mean values since there were only 26 sites with some waste management and the frequency distribution was proportional between the groups.

The site characteristics which showed a negative group relationship are difficult to explain. Some possible explanations are offered here but further research should be conducted to determine the cause of the inverse relationship before major model changes in variable group relationships are implemented.

The characteristic land use in the vicinity of the site showed a negative group relationship. Frequency information was reviewed to determine the source of this inverse relationship. All four rating levels (0-remote, 1-agricultural, 2-commercial/industrial, and 3-residential) occur in each group. The predominant level for the no action group is commercial/industrial with 20 of 32 sites in that level. On the other hand, 50 of the 76 sites in the required action group are in the same level. The major differences appear to be at the extreme rating levels. A greater proportion of sites in 0-remote or 1-agricultural areas are in required action group. Conversely, a greater proportion of sites in 3-residential areas require no action.

A reasonable explanation for this is that the attitude toward sites in remote or agricultural areas where the disposal occurred was probably "Let's dump it out here where it won't hurt anything" with little concern for the potential hazards of contaminant migration. On the other hand, hazardous waste sites in the vicinity of residential areas probably were controlled more because of concern for nearby residents. The land use factor was coded to indicate little adverse impact in remote areas and maximum adverse impact in residential areas. It appears that hazardous waste sites were often located where little adverse impact could occur. The net result was that there was a greater potential for contaminant migration at remote sites. If this is true as the data suggests, then the site location is an indication of potential contaminant migration. It could be used as a multiplier of the pathways and/or waste characteristic category score if the coding was reversed to maintain the same HARM relationship.

Net precipitation is the annual precipitation minus the evapotranspiration. The theory is that net precipitation aids in the migration of contaminants via ground water or surface water. The mean values for net precipitation were 0.6 higher in the no action group, which is the inverse of the general theory. Two possible explanations for this were considered. One is that high annual precipitation would completely flush out the site and remove the hazardous waste. This is not very likely unless the site is very small.

The other possible explanation stems from the sample data. One base had all of the sites that were coded three (high net precipitation). Twelve of the eighteen sites on that base were in the no action group and six were in the required action group. This resulted in a disproportionate number (37.5%) of no action sites coded for high net precipitation. The fact that two-thirds of the sites with high net precipitation required no action may have been due to a unique feature of that base. A fairly large number of sites with low HARM scores was investigated during Phase II. This generally is not done at other bases. Therefore, assuming the HARM score is a good predictor of sites requiring action, a disproportionate share of sites with low HARM scores (predicting no action required) were investigated at a high net precipitation base. This situation could also help explain why the no action group had a higher mean for subsurface flow than the action required group. A high subsurface flow indicates that the bottom of a site is frequently submerged below the ground water level and sites in a high net precipitation area are more likely to have a high ground water level. This is especially true in the coastal area where this base is located.

The other site relationship reviewed dealt with soil and surface permeability. A high soil permeability provides a good pathway for ground water contamination. There was a positive variable relationship, with the mean for soil permeability almost .4 higher in the action required group than the no action group. On the other hand, high impermeability of a surface cover at a site would increase runoff, which in turn would increase chances of surface water contamination. Because of this the surface permeability was coded so that a high impermeability received a high score.

The data shows a slightly higher mean (.16) for the no action group. This small difference is probably not significant but could be attributed to the fact that about 64% of the sites requiring remedial action had high ground water pathway category scores. In those cases the surface permeability score could be either high or low and still not be a determining factor, since the surface water pathway was not the highest pathway.

Phase I Group Profile Differences. Profiles are provided in Tables VI and VII for the Phase I recommended groups to show how HARM was being implemented. A comparison of mean values for each Phase I group was conducted in the same manner as for the Phase II groups. Results are also included in Table V. The relationships were generally similar to those of the Phase II groups. Almost all the variable relationships supported the HARM approach with higher frequency values and means associated with the required action group. This was the expected result because it was a direct application of the HARM theory and only adjusted when judgement influenced site group recommendations.

The waste hazard rating, which showed no mean difference between Phase II groups, had a positive variable relationship in the Phase I groups. The land use mean values also showed a positive variable relationship with the Phase I groups. This would indicate judgement did not perceive the inverse relationship for this variable.

The mean values for subsurface flows went from a negative variable relationship with Phase II groups to a strong positive relationship with Phase I groups. There was a significant shift away from the base which had 12 Phase II no action groups and high subsurface flows to Phase I no action sites with low subsurface flows. This could explain the shift in relationship and results in another positive variable.

Table VI

Profile of Phase I No Action Sites

Site Variable	Mode/%	2nd/%	Code Range	Actual Range
Population	1/47%	0/35%	0-3	0-3
Distance to Well	1/35	3/35	0-3	0-3
Land Use	2/76	0/24	0-3	0-2
Distance within Base	1/53	2/24	0-3	1-3
Critical Environment	0/53	2/24	0-3	0-3
Surface Water Quality	1/65	0/18	0-3	0-2
Ground Water Use	0/29	2/29	0-3	0-3
Population Use of SW	0/100	1/0	0-3	0-0
Population Use of GW	3/65	1/29	0-3	1-3
Waste Quantity	1/59	2/35	1-3	1-3
Confidence Level	1/65	2/35	1-2	1-2
Waste Hazard Rating	3/65	2/24	1-3	1-3
Persistence	1/53	.8/29	.4-1	.4-1
Physical State	1/76	.75/12	.5-1	.5-1
Direct Contamination	0/100	1/0	0-1	0-0
Indirect Contam.	0/94	1/6	0-1	0-1
Distance to SW	3/47	1/24	0-3	0-3
Surface Erosion	0/76	1/12	0-3	0-2
Surface Permeability	1/53	0/47	0-3	0-1
Rainfall Intensity	2/76	3/24	0-3	2-3
Flooding Zone	0/82	1/18	0-3	0-1
Depth to Ground Water	3/53	2/29	0-3	1-3
Net Precipitation	2/41	3/24	0-3	0-3
Soil Permeability	3/59	2/41	0-3	2-3
Subsurface Flows	0/65	1/24	0-3	0-3
Direct Access to GW	0/82	1/18	0-3	0-1
Waste Management	1/88	.95/12	.95-1	.95-1
	Mean	Standard Deviation		
Receptors Score	42.6	11.8	0-100	25-57
Hazardous Waste Score	40.5	17.0	4-100	4-80
Pathways Score	58.4	11.1	0-100	42-80
Final HARM Score	46.9	6.2	0-100	29-55

GW-Ground Water
SW-Surface Water

Table VII

Profile of Phase I Required Investigation Sites

Site Variable	Mode/%	2nd/%	Code Range	Actual Range
Population	0/44%	1/32%	0-3	0-3
Distance to Well	3/55	1/16	0-3	0-3
Land Use	2/63	0/18	0-3	0-3
Distance within Base	3/56	2/35	0-3	0-3
Critical Environment	0/46	2/33	0-3	0-3
Surface Water Quality	1/42	2/24	0-3	0-3
Ground Water Use	0/44	1/27	0-3	0-3
Population Use of SW	0/86	3/9	0-3	0-3
Population Use of GW	3/69	1/13	0-3	0-3
Waste Quantity	1/36	2/34	1-3	1-3
Confidence Level	2/54	1/46	1-2	1-2
Waste Hazard Rating	3/90	2/10	1-3	2-3
Persistence	1/75	.9/15	.4-1	.8-1
Physical State	1/84	.75/15	.5-1	.5-1
Direct Contamination	0/92	1/8	0-1	0-1
Indirect Contam.	0/78	1/22	0-1	0-1
Distance to SW	3/63	2/19	0-3	0-3
Surface Erosion	0/51	1/34	0-3	0-3
Surface Permeability	0/46	1/46	0-3	0-3
Rainfall Intensity	2/44	3/36	0-3	1-3
Flooding Zone	0/69	1/24	0-3	0-3
Depth to Ground Water	3/60	2/27	0-3	1-3
Net Precipitation	2/43	0/21	0-3	0-3
Soil Permeability	3/46	2/34	0-3	1-3
Subsurface Flows	1/30	0/30	0-3	0-3
Direct Access to GW	0/65	3/24	0-3	0-3
Waste Management	1/74	.95/26	.95-1	.95-1
	Mean	Standard Deviation		
Receptors Score	49.4	15.5	0-100	28-91
Hazardous Waste Score	60.2	21.3	4-100	23-100
Pathways Score	74.2	16.5	0-100	28-100
Final HARM Score	60.4	10.0	0-100	44-85

GW-Ground Water
SW-Surface Water

The two Phase I group variables with mean values that indicate a negative relationship are the net precipitation and soil permeability. The mean differences were not that large and it was difficult to identify a cause. It could possibly be attributable to the fact that only 17 Phase I no action sites may not have been a large enough sample for proper statistical analysis.

HARM Effectiveness

If the HARM model could predict with total accuracy there would be no need to confirm the results with investigative Phase II studies. This is not a realistic expectation, so some errors will normally occur. Ideally the model should minimize all classification errors, especially those errors predicting no action at sites that actually required action.

The first research objective was to validate the current HARM ranking system by evaluating HARM's ability to correctly classify sites. Two measures of effectiveness were used throughout this research. One measurement was that of the overall percentage of sites correctly classified. The other was the number of sites incorrectly predicted to require no action when action was actually required. For the purpose of this research this type error will be referred to as a Class I error. This second measurement was singled out because it is considered the more serious type of error. If an error is to be made it is desirable to err on the side of safety. In other words, it is better to investigate a few sites that do not need investigation than to not investigate sites that require investigation.

The results for the current HARM model as validated with Phase II classifications are provided in Table VIII. The current HARM model correctly classified 67.6% of the sites into the correct group (required action/no action). This table shows how many sites were predicted for each group as well as how many of those predictions were correct and how many were incorrect.

Table VIII

Classification Results for Phase I Predictions

Actual Group	No. of Cases	Predicted Group Membership: No Action	Predicted Group Membership: Reqd. Action
No Action	32	7 21.9%	25 78.1%
Required Action	76	10 13.2%	66 86.8%

Total Percent of Cases Correctly Classified = 67.6%

Only seventeen sites were predicted to require no action, but ten of those sites actually required remedial action (Class I error). If the Phase I recommendations had been followed these ten sites would not have been investigated, although the sites were actually contaminated. On the other hand, twenty-five sites were investigated which did not actually have a problem which caused unnecessary investigative expenses.

HARM Improvement

Evaluation of the HARM Model. To improve HARM it was necessary to investigate the components of the current HARM model. The current procedure is to use the final HARM score and some independent judgement to predict if the site needs further investigation (required action) or not (no action). That is why there is some final HARM score overlap between groups as shown in the Phase I group profiles, Tables VI and VII. The highest HARM score in the group recommended for no action was fifty-five, while the lowest HARM score in the Phase I required action group was forty-four. This exercise of judgement instead of just using a cut-off score is supported by the actual Phase II group profiles in Tables III and IV. The lowest value in the Phase II action required group was forty-three while the highest value in the Phase II no action group was seventy-one.

The discriminant analysis program was run with only the final HARM score as the predictor variable to evaluate the use of judgement against a single cut-off score. The results show several important things. The percentage of sites correctly classified was 64.8%, which was only slightly less than when judgement was used, but the number of Class I errors was thirty, which was significantly worse. The canonical correlation for the discriminant function based on the total HARM score was 0.387. This would indicate that there is some relationship, but it is not strong.

The use of a single cut-off score resulted in 54 cases being assigned to the no action group and 46 cases being assigned to the required action group. When judgement was applied, the figures were 17 and 91 respectively. Clearly, the exercise of judgement resulted in more cases being assigned to the required action group. This bias significantly reduced the Class I error at the expense of investigating more sites than necessary. This fact both emphasizes the importance attached to committing a Class I error and the need for a better classification method to reduce excessive cost of unnecessary investigations. The results of this run are presented in Table IX.

Table IX

Classification Results for Final HARM Scores

Actual Group	No.of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	24 75.0%	8 25.0%
Required Action	76	30 39.5%	46 60.5%

Total Percent of Cases Correctly Classified = 64.8%
Canonical Correlation = .378

One other statistic was checked to insure that the discriminant function reflected actual group differences.

The chi-squared significance test for this run resulted in a P-value of .0001. This meant that the chances of finding a discriminant function this good, when there was actually no difference between groups, was one in 10,000. Thus, we can assume the discriminant function represents actual group differences. For all subsequent analysis, the P-value was not reported unless it was greater than .001.

The next approach was to use HARM category scores as predictor variables, and to evaluate their contribution to the discriminant function's ability to correctly classify groups. Three major category scores (receptors, waste characteristics, and pathways) were multiplied by the management factor just like current HARM procedures, and then used as predictor variables. The results presented in Table X show some overall improvement in classification (69.4%) but little improvement in the Class I errors (29). The canonical correlation also improved to .421.

Table X

Classification Results for Three HARM Categories

Actual Group	No. of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	28 87.5%	4 12.5%
Required Action	76	29 38.2%	47 61.8%

Total Percent of Cases Correctly Classified = 69.4%
Canonical Correlation = .421

The HARM model considers the waste management factor as a separate category, although it is used as a multiplier. The next approach, therefore, was to use all four category scores as predictor variables. The management factor and the other three unadjusted category scores provided the best overall results at this point and are presented in Table XI. The overall classification was 68.5%, which was not quite as good as with the three adjusted category scores. However, the canonical correlation was .476 and the Class I errors dropped to twenty-five. These results were starting to show some improvement over the current HARM model.

Table XI

Classification Results for Four Categories

Actual Group	No. of Cases	Predicted Group Membership	
		No Action	Reqd. Action
No Action	32	23 71.9%	9 28.1%
Required Action	76	25 32.9%	51 67.1%

Total Percent of Cases Correctly Classified = 68.5%
Canonical Correlation = .476

At this point the model improvement approach was changed. Discriminant analysis was run with all 27 HARM predictor variables. This approach used all the HARM information but did not consider any special HARM model interaction or structural features. The results of this approach are provided in Table XII.

There was marked improvement in all areas. The canonical correlation jumped to .740, which would indicate a fairly strong relationship. The percentage of all sites correctly classified increased to 88.9% and the number of Class I errors dropped to nine.

Table XII

Classification Results for All HARM Variables

Actual Group	No.of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	29 90.6%	3 9.4%
Required Action	76	9 11.8%	67 88.2%

Total Percent of Cases Correctly Classified = 88.9%
Canonical Correlation = .740

To see if any special features of the HARM model algorithm could improve on the basic linear model approach, several special predictor variables were considered. The two areas of the model that use special features are in the waste characteristics and pathways categories.

Each special feature category score was used in place of the factors they represented, but in both cases the predictive ability of the discriminant function decreased. Although these special HARM features did not improve the discriminant function, other special predictor variables will be tried in the next section.

Interaction and Structural Changes. The best discriminant function developed at this point contains no interaction predictor variables. The HARM model considered interaction between the waste characteristic category score and the waste management factor. This interaction appeared logical. Therefore, a similar interaction variable was developed. The key factor was the waste quantity, which the descriptive statistics showed was an important factor. This factor was multiplied by the waste persistence factor and the waste management factor to form a new interaction predictor variable. This combination of factors was used because it was felt that their combined effects could not be represented by the linear combination of individual factors.

The new interaction predictor variable was added to the list of basic HARM factors which were already predictor variables in the discriminant function. The results shown in Table XIII indicate some minor improvement. The canonical correlation increased from .740 to .750, Class I errors dropped from 9 to 8, and the overall correct classification was 89.8%.

The descriptive statistics indicated that the land use factor may be a good indicator of hazardous waste management if coded in reverse. This was done and the recoded land use factor was substituted for the waste management factor in

Table XIII

Classification Results for Hazardous Waste Interaction

Actual Group	No. of Cases	Predicted Group Membership: No Action	Predicted Group Membership: Reqd. Action
No Action	32	29 90.6%	3 9.4%
Required Action	76	8 10.5%	68 89.5%

Total Percent of Cases Correctly Classified = 89.8%
Canonical Correlation = .750

the interaction predictor variable. The results shown in Table XIV indicated a slight decrease in the discriminant function's ability to classify sites. The land use factor apparently was an unsatisfactory indicator of waste management.

Table XIV

Classification Results for Interaction With Land Use

Actual Group	No. of Cases	Predicted Group Membership: No Action	Predicted Group Membership: Reqd. Action
No Action	32	29 90.6%	3 9.4%
Required Action	76	10 13.2%	66 86.8%

Total Percent of Cases Correctly Classified = 88.0%
Canonical Correlation = .740

Another area of possible interaction dealt with matching receptor factors with pathway factors. Both the ground water and surface water pathways had receptors associated with them. The concept was that the combined effect of receptors and pathways would be greater than the effect of their individual factors. The ground water interaction was investigated first because ground water was used more by the surrounding population. In addition, two-thirds of the sites requiring remedial action had ground water as the most probable pathway.

The ground water use factor was added to the population use of ground water factor. This composite receptor variable was then multiplied by the sum of the pathway factors; depth to ground water, soil permeability and direct access to ground water. The subsurface flow factor was excluded from the pathway factors because it showed an inverse relationship with the other pathway factors. The resultant special variable was considered the best combination to represent this interaction.

The results of adding this interaction variable to the basic HARM factors already in the discriminant function, as shown in Table XV, did not indicate any model improvement. Since ground water interaction showed no improvement, surface water interaction was not tested. This approach was considered reasonable because 88% of all sites had no use of the surface water by the surrounding population.

The next approach was to try adding a predictor variable that represented the highest pathway subscore.

Table XV

Classification Results for Ground Water Interaction

Actual Group	No. of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	29 90.6%	3 9.4%
Required Action	76	9 11.8%	67 88.2%

Total Percent of Cases Correctly Classified = 88.9%
Canonical Correlation = .740

This added emphasis to the highest individual pathway. The approach essentially duplicated the HARM procedure of selecting the highest subscore among the five subcategories (direct evidence, indirect evidence, surface water, flooding, and ground water). This new predictor variable was added to the list of HARM factor predictor variables. As with the past run, adding this special variable to the discriminant function did not improve its classification capability. This is what the results in Table XVI indicate, although the canonical correlation increased slightly to .758.

A similar structural feature was tried by creating a variable that recorded the highest pathway subscore between just the ground water and the surface water. This approach was taken to add emphasis to the highest score among the two most common pathways. Again the results showed no improvement.

Table XVI

Classification Results for Maximum Pathway

Actual Group	No. of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	29 90.6%	3 9.4%
Required Action	76	9 11.8%	67 88.2%

Total Percent of Cases Correctly Classified = 88.9%
Canonical Correlation = .758

This concluded the search for special interaction and structural variables. Although only the waste quantity/waste persistence/management factor interaction variable showed overall improvement, the other special variables developed were also retained as possible predictor variables. The main reason they were kept was because they represented important concepts which might become significant in later analysis.

Additional Site Information. During the collection of data for this research several site characteristics were recorded which were not being used by the HARM model. Information on the type of hazardous waste site, the type of probable contamination, and whether or not the disposal was on or below ground surface was collected.

Stepwise discriminant analysis was used to select only the additional variables with the most discriminating power. The objective was to add only additional variables that significantly added to the discriminant functions ability to classify sites. The list of variables used included all the HARM factor variables, the special interaction and structural variables, and the new information variables. All variables were allowed to enter the stepwise analysis. Several runs were required to give all the new information variables a chance to enter the stepwise analysis.

The results represent the maximum improvement that was obtained from this research using all the information available and adding special interaction and structural variables. The results of this discriminant function model are presented in Table XVII.

Table XVII

Classification Results for Full Information Model

Actual Group	No.of Cases	Predicted Group Membership No Action	Predicted Group Membership Reqd. Action
No Action	32	30 93.8%	2 6.3%
Required Action	76	5 6.6%	71 93.4%

Total Percent of Cases Correctly Classified = 93.5%
Canonical Correlation = .838

The results show that the additional information improved the model's ability to correctly classify sites to 93.5%, with only five sites in the Class I error category. A canonical correlation of 0.838 also indicated a strong relationship. This is the final full information model

Model Simplification

The collection of information can be time consuming and expensive, so an attempt was made to reduce the number of predictor variables necessary to get good classification results. Again, stepwise analysis was used to select the best predictor variables. The HARM factor variables, the special variables, and the significant new information variables were all allowed to enter the stepwise analysis. This approach eliminated fourteen predictor variables, but the percentage of correctly classified sites dropped to 90.7% with 9 Class I errors and a canonical correlation of 0.815. Most of the information contained in the site characteristic variables eliminated was provided by other variables with similar information or was of little significance. Table XVIII summarizes the results of this reduced model.

Table XVIII

Classification Results for Reduced Model

Actual Group	No.of Cases	Predicted Group Membership No Action	Reqd. Action
No Action	32	31 96.9%	1 3.1%
Required Action	76	9 11.8%	67 88.2%

Total Percent of Cases Correctly Classified = 90.7%
Canonical Correlation = .815

Interpretation of Model

The discriminant function coefficients obtained from this research analysis can be used to indicate the importance of the site predictor variables. The absolute value of the weights which are the standardized canonical discriminant function coefficients, is an indication of the importance of the variables. For example, in the full model the waste quantity with a coefficient of 2.88 contributes about three times as much discriminating ability as the waste persistence factor with a coefficient of .89. If the investigator makes an evaluation error on a heavily weighted variable it could easily result in a classification error. Therefore, extra care should be taken when evaluating the heavily weighted variables. The list of all the variables and their associated weights are provided in Table XIX.

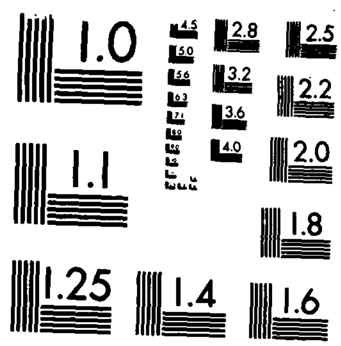
AD-A161 066 VALIDATION OF AIR FORCE HAZARD ASSESSMENT RATING METHODOLOGY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYSTEMS AND LOGISTICS M C ANDERSON 2/2

UNCLASSIFIED SEP 85 AFIT/GEM/LSY/85S-1 F/G 6/20 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Table XIX

Standardized Discriminant Function Coefficients

Site Variable	Full Model	Reduced Model
Population	0.176	-0.252
Distance to Well	0.686	-0.561
Land Use	-0.210	0.314
Distance within Base	-0.209	
Critical Environment	-0.289	
Surface Water Quality	-0.268	0.314
Ground Water Use	-1.363	1.127
Population Use of SW	1.059	-0.993
Population Use of GW	-1.097	0.934
Waste Quantity	2.880	-0.343
Confidence Level	-0.161	0.207
Waste Hazard Rating	0.888	
Persistence	-0.208	
Physical State	0.063	
Direct Contamination	0.261	
Indirect Contamination	0.463	-0.280
Distance to Surface Water	0.417	-0.258
Surface Erosion	0.028	
Surface Permeability	0.574	
Rainfall Intensity	-0.076	
Flooding Zone	-0.125	0.270
Depth to Ground Water	0.454	-0.485
Net Precipitation	-1.779	1.702
Soil Permeability	1.352	-0.607
Subsurface Flows	-0.278	0.462
Direct Access to GW	0.371	
Waste Management	0.404	-0.380
Landfill Site	1.320	-1.110
Fire Training Area	0.160	
Spill Site	-0.335	0.279
Oil/Water Separator Area	-0.286	0.240
Below Ground Disposal	-0.598	0.707
Radiological Waste Site	-0.181	0.322
Maximum Pathway	0.254	-0.550
Waste/management Interaction	-2.733	
GW/Receptor Interaction	-0.077	
Maximum GW/SW Pathway	-0.196	

GW-Ground Water
SW-Surface Water

Classification is the process of deciding to which group (no action, required action) a site belongs. The HARM model provided a final HARM score which was used to decide to which group a site belonged. The discriminant analysis program derives a classification function for each group from the discriminant function. Each site has a classification score computed for each group. The classification score is obtained by multiplying each site predictor variable score by its associated classification function coefficient and summing all the results plus a function constant. The classification function can be expressed by the following equation (20:43):

$$C_k = B_{ko} + B_{ki}V_i + \ldots + B_{kp}V_p$$

where

C_k = the classification score for group k

V_i = the value of the ith predictor variable

B_{ko} = classification function constant for group k

B_{ki} = ith classification coefficient for the ith predictor variable in group k

k = the number of the group

i = the number of the variable

p = the total number of predictor variables

The site is assigned to the group for which it receives the highest classification score. The classification function coefficients provided in Table XXI can now be used to classify sites instead of the HARM model. Both the full information model and the reduced information model coefficients are provided, so either can be used.

The list of misclassified sites provided in Table XX were reviewed to check for any unique features. The highest probability column reflects the probability for the group that was selected. Three of the seven misclassified sites had high probabilities for the wrong group so they were studied in more detail.

Table XX

Classification Probability

Site No.	Actual Group	Group Selected	Highest Probability
84	1	0	.5344
87	1	0	.7160
89	0	1	.8154
91	0	1	.5627
92	1	0	.5714
100	1	0	.7523
105	1	0	.5942

0 --No Action

Table XXI

Classification Function Coefficients

Site Variable	Full No Action	Full Action	Reduced No Action	Reduced Action
Population	3.6345	4.2307	4.0588	4.8426
Distance to Well	-19.653	-17.634	-11.718	-10.207
Land Use	-34.734	-35.538	-31.724	-32.524
Distance within Base	15.236	14.300		
Critical Environment	-10.443	-11.282		
Surface Water Quality	94.229	93.256	57.842	56.800
Ground Water Use	93.075	89.252	15.704	12.814
Population Use of SW	39.957	44.373	0.5428	4.3325
Population Use of GW	173.66	170.31	41.150	38.539
Waste Quantity	361.68	374.04	31.777	33.126
Confidence Level	10.208	9.1309	5.3782	4.1158
Waste Hazard Rating	230.82	237.94		
Persistence	-367.66	-375.50		
Physical State	69.940	71.721		
Direct Contamination	-158.52	-154.67		
Indirect Contamination	-62.195	-58.238	-15.389	-13.198
Distance to Surface Water	1.9304	3.5686	-14.927	-13.999
Surface Erosion	32.694	32.800		
Surface Permeability	88.523	91.514		
Rainfall Intensity	-30.341	-30.702		
Flooding Zone	-52.860	-53.428	-26.828	-27.953
Depth to Ground Water	74.849	76.953	15.815	17.872
Net Precipitation	-3.5711	-9.7068	10.355	4.9849
Soil Permeability	132.38	138.56	51.500	-48.960
Subsurface Flows	19.455	18.619	-6.9267	-8.1975
Direct Access to GW	90.285	91.303		
Waste Management	7849.9	7912.3	5367.4	5921.2
Landfill Site	-0.2493	8.7317	-17.212	-10.303
Fire Training Area	28.706	30.415		
Spill Site	-34.065	-38.621	-6.4401	-9.9049
Oil/Water Separator Area	154.96	145.00	152.01	145.34
Radiological Waste Site	187.04	182.51	98.473	91.083
Below Ground Disposal	58.725	54.756	58.230	53.938
Maximum Pathway	39.117	39.430	23.603	24.224
Waste/Mgt. Interaction	-113.78	-117.69		
GW/Receptor Interaction	-20.975	-20.997		
Maximum GW/SW Pathway	-51.620	-52.105		
Constant	-4504.5	-4584.3	-3045.0	-3099.1

GW-Ground Water
SW-Surface Water

No unique variables were identified that affected all three sites. Instead, the prediction errors seemed to be from a combination of variables. The variables with discriminant function coefficients greater than one were looked at closely because different values for these variables can significantly affect the group prediction for the site.

The site predicted to require action when none was required, had several variable values that were inconsistent with other values in that group. The site was classified as a landfill disposal type. Fifty-one percent of the required action group were of the landfill disposal type, while only twenty-eight percent of the sites in the no action group were landfills. Also, this site had a soil permeability value of two, and eighty-eight percent of the required action group was coded two or higher. On the other hand, seventy-two percent of the no action group were coded two or higher. In addition the population near the site was coded two and only nine percent of the no action sites were coded two or more, while twenty-nine percent of the action required group was in this range. This shows the values support the required action group prediction and partially explain why the wrong group was selected.

The other two sites were predicted to require no action when action was necessary (Class I error). One of the sites was classified as a landfill, which does not support the prediction of no action. One variable identified that supported the no action classification for the landfill site was the distance within the base variable which was coded one. The required action group only had ten percent of its sites coded less than two while the no action group had twenty-eight percent of its sites coded less than two. The other site was not a landfill, which supported the no action prediction.

The only unique feature that could be identified was that all seven of the misclassified sites were at two of the first bases investigated during Phase I. There is normally a learning curve when a new evaluation procedure is started and sometimes initial evaluations are inconsistent. This could help account for the misclassifications if there were incorrect evaluations of site variables.

Model Accuracy

To check the accuracy of the classification functions, approximately twenty percent of the sites in each group was deleted from the data and new classification functions were developed. The larger group of sites was used to develop the classification functions to help insure the stability of the function coefficients. Then the ability of the new classification functions to classify the excluded sites was evaluated by looking at the classification results. Although the remaining group of sites was probably too small of a sample to accurately test the classification ability, the results of three runs are provided in Table XXII. The results show improvement over the original HARM model in all cases. Although the overall classification showed a wide range, the percentage correctly classified in each of the correct groups was greater than 80% for five of the six groups.

The canonical correlation stayed above .8 for all three discriminant functions run with the split samples. The ability of the new models to classify the remaining sites ranged from seventy-five percent to ninety-two percent in overall classification efficiency.

Table XXII

Classification Results for Excluded Sites

Actual Group	No. of Cases	Predicted Group Membership No Action	Reqd. Action
No Action	7	4 57.1%	3 42.9%
Required Action	17	3 17.6%	14 82.4%
Total Percent of Cases Correctly Classified = 75.0% Canonical Correlation with Included Sites = .883			
No Action	8	8 100.0%	0 0.0%
Required Action	19	2 10.5%	17 89.5%
Total Percent of Cases Correctly Classified = 92.6% Canonical Correlation with Included Sites = .833			
No Action	7	7 100.0%	0 0.0%
Required Action	17	3 17.6%	14 82.4%
Total Percent of Cases Correctly Classified = 87.5% Canonical Correlation with Included Sites = .822			

VI. Conclusions and Recommendations

The overall goal of this research was to improve the ability to correctly predict which hazardous waste sites needed further investigation based on an initial Phase I Records Search. This goal was achieved and important knowledge about site characteristics and the HARM model was gained. Conclusions about the research questions outlined in Chapter I will be presented first, along with some general conclusions about the research data. The last section will provide some final recommendations.

Conclusions

Research Question One. What does IRP Phase II data show about HARM's ability to predict hazardous waste sites that need investigation?

HARM's predictive capability both with and without judgement was documented. By using the final HARM score and expert judgement 67.6% of the sites were predicted to belong to the correct groups. Ten sites were recommended for no action that should have required action based on Phase II results. When judgement was eliminated from the final classification the percentage of correctly classified sites dropped to 64.8%, but 30 sites were recommended for no action when they should have required action based upon Phase II results. This was a marked decrease in the effectiveness of the HARM model and reflected the the results of conservative judgement. The HARM model was not very effective in predicting the correct group, but with judgement it was significantly better than randomly assigning sites to groups.

Research Question Two. Can HARM's predictive capability be improved?

The answer to this question was a definite "yes". By using discriminant analysis to properly weight site characteristics and develop classification functions 93.5%, of the sites were properly classified with only five sites predicted for no action when action was required. The canonical correlation coefficient was .838, which indicated good association between the groups and the discriminant function. The site characteristic variables used to obtain these results included twenty-seven raw HARM factors, two structural and two interaction variables, and six new site characteristic variables.

The accuracy of the full model was evaluated using a split-sample approach. The results were inconclusive because of the small sample size but the approach still showed improvement over HARM's original classification efficiency.

Research Question Three. Can HARM maintain its effectiveness in predicting contaminated sites with less information?

Once the full model was developed with all available information, stepwise analysis was used to obtain the reduced model which had 14 less variables. With 24 predictor variables the reduced model was still very effective, correctly classifying 90.7% of the sites. The canonical correlation remained high (.815) but the number of Class I errors went from five to nine. Although not as good as the full model, the overall results were still considerably better than the original HARM model.

The results show that good overall classification ability can still be maintained even with reduced information. The increase in Class I errors is not desirable, however. Therefore the full model is recommended for use whenever possible.

Research Question Four. Does HARM work equally well for all types of sites and situations?

No unique types of sites or situations were identified that could account for the misclassification of sites. It was noted, however, that all the misclassified sites were at bases where the HARM model was first used. This indicated there was probably a learning curve associated with using the model. Improper variable coding and resultant classification errors could be minimized by only allowing experienced people to use the model.

The data used for the research was considered reliable and accurate. Descriptive statistics did reveal some surprising relationships between the Phase II classification groups and the site characteristic variables. Some variables strongly supported current HARM model relationships, but three variables appeared to have strong inverse relationships to the current theory. The three variables were: the surrounding land use factor from the receptor category, net precipitation, and subsurface flows. Some explanations were offered for the lower variable values in the required action group, but more research is needed to evaluate these inverse relationships.

Recommendations

This research study showed that the current HARM model is doing a reasonable job in predicting which sites need further investigation, but there is room for improvement. Both the full and reduced model developed with discriminant analysis showed over a 20 percentage point increase in predictive ability. Although the results with the new models are very good, they need further testing and evaluation before they are ready for general use.

There is one additional recommendation. Currently, there is no bias in the model to err on the side of safety. Each site has an equal chance of being assigned to either group. By multiplying the group classification scores by a probability proportionate to the desired safety factor, a more conservative classification process can be developed.

A 60% probability for the required action group, and a 40% probability for the no action group should provide sufficient safety bias. The result will be that marginal sites which could easily be assigned to either group will be assigned to the required action group.

Future Research

With a significant improvement in the model's predictive capability demonstrated as achievable, the final evaluation and testing necessary before general application is warranted. This study was limited primarily by the number of sites in the no action group. Although technically sufficient for the study, the sites may not have constituted a representative sample. This possibility was highlighted by the unexpected inverse relationship with net precipitation. Twelve of the no action sites were from one high net precipitation base and 20 of the 32 no action sites were from just two bases.

Increasing the number and variety of no action sites will insure a more representative sample. It will also help in the investigation of the inverse relationships which must be understood before any model incorporating them can be used. All the model variable relationships should have plausable explanations before they are included. Once additional data is obtained, a final model can be developed by using the same research approach used here.

One area for future research which was not considered is the coding criteria used for the site characteristics. This may be especially beneficial for the heavily weighted variables. In any case the refinement and testing of the models developed in this research appears to be justified by the potential benefits.

APPENDIX A: USAF INSTALLATION RESTORATION PROGRAM HAZARD ASSESSMENT RATING METHODOLOGY

Extracted from (18)

BACKGROUND

The Department of Defense (DoD) has established a comprehensive program to identify, evaluate, and control problems associated with past disposal practices at DoD facilities. One of the actions required under this program is to:

> "develop and maintain a priority listing of contaminated installations and facilities for remedial action based on potential hazard to public health, welfare, and environmental impacts." (Reference: DEQPPM 81-5, 11 December 1981).

Accordingly, the United States Air Force (USAF) has sought to establish a system to set priorities for taking further actions at sites based upon information gathered during the Records Search phase of its Installation Restoration Program (IRP).

The first site rating model was developed in June 1981 at a meeting with representatives from the USAF Occupational and Environmental Health Laboratory (OEHL), Air Force Engineering and Services Center (AFESC), Engineering-Science (ES) and CH2M HILL. The basis for this model was a system developed for EPA by JRB Associates of McLean, Virginia. The JRB model was modified to meet Air Force needs.

After using this model for 6 months at over 20 Air Force installations, certain inadequacies became apparent. Therefore, on January 26 and 27, 1982, representatives of

USAF OEHL, AFESC, various major commands, Engineering Science, and CH2M HILL met to address the inadequacies. The result of the meeting was a new site rating model designed to present a better picture of the hazards posed by sites at Air Force installations. The new rating model described in this presentation is referred to as the Hazard Assessment Rating Methodology.

PURPOSE

The purpose of the site rating model is to provide a relative ranking of sites of suspected contamination from hazardous substances. This model will assist the Air Force in setting priorities for follow-on site investigations and confirmation work under Phase II of IRP.

This rating system is used only after it has been determined that (1) potential for contamination exists (hazardous wastes present in sufficient quantity), and (2) potential for migration exists. A site can be deleted from consideration for rating on either basis.

DESCRIPTION OF MODEL

Like the other hazardous waste site ranking models, the U.S. Air Force's site rating model uses a scoring system to rank sites for priority attention. However, in developing this model, the designers incorporated some special features to meet specific DoD program needs.

The model uses data readily obtained during the Record Search portion (Phase I) of the IRP. Scoring judgments and computations are easily made. In assessing the hazards at a given site, the model develops a score based on the most likely routes of contamination and the worst hazards at the site. Sites are given low scores only if there are clearly no hazards at the site. This approach meshes well with the policy for evaluating and setting restrictions on excess DoD properties.

Site scores are developed using the appropriate ranking factors according to the method presented in the flow chart (Figure H-1). The site rating form is provided on Figure H-2 and the rating factor guidelines are provided in Table H-1.

As with the previous model, this model considers four aspects of the hazard posed by a specific site: the possible receptors of the contamination, the waste and its characteristics, the potential pathways for waste contaminant migration, and any efforts to contain the contamination. Each of these categories contains a number of rating factors that are used in the overall hazard rating.

The receptors category rating is calculated by scoring each factor, multiplying by a factor weighting constant, and adding the weighted scores to obtain a total category score.

The pathways category rating is based on evidence of contaminant migration or an evaluation of the highest potential (worst case) for contaminant migration along one of three pathways. If evidence of contaminant migration exists, the category is given a subscore of 80 to 100 points. For indirect evidence, 80 points are assigned and for direct evidence 100 points are assigned. If no evidence is found, the highest score among three possible routes is used. These routes are surface-water migration, flooding, and ground-water migration. Evaluation of each route involves factors associated with the particular migration route. The three pathways are evaluated and the highest score among all four of the potential scores is used.

The waste characteristics category is scored in three steps. First, a point rating is assigned based on an assessment of the waste quantity and the hazard (worst case) associated with the site. The level of confidence in the information is also factored into the assessment. Next, the score is multiplied by a waste persistence factor, which acts to reduce the score if the waste is not very persistent. Finally, the score is further modified by the physical state of the waste. Liquid wastes receive the maximum score, while scores for sludges and solids are reduced.

The scores for each of the three categories are then added together and normalized to a maximum possible score of 100. Then the waste management practice category is scored. Scores for sites at which there is no containment are not reduced. Scores for sites with limited containment can be reduced by 5 percent. If a site is contained and well managed, its score can be reduced by 90 percent. The final site score is calculated by applying the waste management practices category factor to the sum of the scores for the other three categories.

Table XXIII (H-1)

HAZARDOUS ASSESSMENT RATING METHODOLOGY GUIDELINES

I. **RECEPTORS CATEGORY**

Rating Factors	Rating Scale Levels 0	1	2	3	Multiplier
A. Population within 1,000 feet (includes on-base facilities)	0	1-25	26-100	Greater than 100	4
B. Distance to nearest water well	Greater than 3 miles	1 to 3 miles	3,001 feet to 1 mile	0 to 3,000 feet	10
C. Land Use/Zoning (within 1-mile radius)	Completely remote (zoning not applicable)	Agricultural	Commercial or Industrial	Residential	3
D. Distance to installation boundary	Greater than 2 miles	1 to 2 miles	1,001 feet to 1 mile	0 to 1,000 feet	6
E. Critical environments (within 1-mile radius)	Not a critical environment	Natural areas	Pristine natural areas; minor wetlands; preserved areas; presence of economically important natural resources susceptible to contamination	Major habitat of an endangered or threatened species; presence of recharge area; major wetlands	10
F. Water quality/use designation of nearest surface water body	Agricultural or Industrial use	Recreation, propagation and management of fish and wildlife	Shellfish propagation and harvesting	Potable water supplies	6
G. Ground-water use of uppermost aquifer	Not used, other sources readily available	Commercial, industrial, or irrigation, very limited other water sources	Drinking water, municipal water available	Drinking water, no municipal water available; commercial, industrial, or irrigation, no other water source available	9
H. Population served by surface water supplies within 3 miles downstream of site	0	1-15	51-1,000	Greater than 1,000	6
I. Population served by aquifer supplies within 3 miles of site	0	1-50	51-1,000	Greater than 1,000	6

Table XXIII (H-1) Continued

II. WASTE CHARACTERISTICS

A-1 Hazardous Waste Quantity

S = Small quantity (5 tons or 20 drums of liquid)
M = Moderate quantity (5 to 20 tons or 21 to 85 drums of liquid)
L = Large quantity (20 tons or 85 drums of liquid)

A-2 Confidence Level of Information

C = Confirmed confidence level (minimum criteria below)

- o Verbal reports from interviewer (at least 2) or written information from the records
- o Knowledge of types and quantities of wastes generated by shops and other areas on base

S = Suspected confidence level

- o No verbal reports or conflicting verbal reports and no written information from the records
- o Logic based on a knowledge of the types and quantities of hazardous wastes generated at the base, and a history of past waste disposal practices indicate that these wastes were disposed of at a site

A-3 Hazard Rating

Rating Factors	Rating Scale Levels 0	1	2	3
Toxicity	Sax's Level 0	Sax's Level 1	Sax's Level 2	Sax's Level 3
Ignitability	Flash point greater than 200°F	Flash point at 140°F to 200°F	Flash point at 80°F to 140°F	Flash point less than 80°F
Radioactivity	At or below background levels	1 to 3 times background levels	3 to 5 times background levels	Over 5 times background levels

Use the highest individual rating based on toxicity, ignitability and radioactivity and determine the hazard rating.

Hazard Rating	Points
High (H)	3
Medium (M)	2
Low (L)	1

Table XXIII (H-1) Continued

II. **WASTE CHARACTERISTICS**--Continued

Waste Characteristics Matrix

Point Rating	Hazardous Waste Quantity	Confidence Level of Information	Hazard Rating
100	L	C	H
80	L	C	M
	M	C	H
70	L	S	H
60	S	C	H
	M	C	M
50	L	S	M
	L	C	L
	M	S	H
	S	C	M
40	S	S	H
	M	S	M
	M	C	L
	L	S	L
30	S	C	L
	M	S	L
	S	S	M
20	S	S	L

Notes:
For a site with more than one hazardous waste, the waste quantities may be added using the following rules:
<u>Confidence Level</u>
- o Confirmed confidence levels (C) can be added.
- o Suspected confidence levels (S) can be added.
- o Confirmed confidence levels cannot be added with suspected confidence levels.

<u>Waste Hazard Rating</u>
- o Wastes with the same hazard rating can be added.
- o Wastes with different hazard ratings can only be added in a downgrade mode, e.g., MCM + SCM = LCM if the total quantity is greater than 20 tons.

<u>Example:</u> Several wastes may be present at a site, each having an MCM designation (60 points). By adding the quantities of each waste, the designation may change to LCM (80 points). In this case, the correct point rating for the waste is 80.

B. **Persistence Multiplier for Point Rating**

Persistence Criteria	Multiply Point Rating From Part A by the Following
Metals, polycyclic compounds, and halogenated hydrocarbons	1.0
Substituted and other ring compounds	0.9
Straight chain hydrocarbons	0.8
Easily biodegradable compounds	0.4

C. **Physical State Multiplier**

Physical State	Multiply Point Total From Parts A and B by the Following
Liquid	1.0
Sludge	0.75
Solid	0.50

Table XXIII (H-1) Continued

III. PATHWAYS CATEGORY

A. Evidence of Contamination

Direct evidence is obtained from laboratory analyses of hazardous contaminants present above natural background levels in surface water, ground water, or air. Evidence should confirm that the source of contamination is the site being evaluated.

Indirect evidence might be from visual observation (i.e., leachate), vegetation stress, sludge deposits, presence of taste and odors in drinking water, or reported discharges that cannot be directly confirmed as resulting from the site, but the site is greatly suspected of being a source of contamination.

B-1 Potential for Surface Water Contamination

Rating Factors	Rating Scale Levels 0	1	2	3	Multiplier
Distance to nearest surface water (includes drainage ditches and storm sewers	Greater than 1 mile	2,001 feet to 1 mile	501 feet to 2,000 feet	0 to 500 feet	8
Net precipitation	Less than -10 inches	-10 to +5 inches	+5 to +20 inches	Greater than +20 inches	6
Surface erosion	None	Slight	Moderate	Severe	8
Surface permeability	0% to 15% clay ($>10^{-2}$ cm/sec)	15% to 30% clay (10^{-2} to 10^{-4} cm/sec)	30% to 50% clay (10^{-4} to 10^{-6} cm/sec)	Greater than 50% clay ($>10^{-6}$ cm/sec)	6
Rainfall intensity based on 1-year 24-hour rainfall	≤1.0 inch	1.0 to 2.0 inches	2.1 to 3.0 inches	≥3.0 inches	8

B-2 Potential for Flooding

Rating Factors	0	1	2	3	Multiplier
Floodplain	Beyond 100-year floodplain	In 100-year floodplain	In 10-year floodplain	Floods annually	1

B-3 Potential for Ground-Water Contamination

Rating Factors	0	1	2	3	Multiplier
Depth to ground water	Greater than 500 feet	50 to 500 feet	11 to 50 feet	0 to 10 feet	8
Net precipitation	Less than -10 inches	-10 to +5 inches	+5 to + 20 inches	Greater than +20 inches	6
Soil permeability	Greater than 50% clay ($\leq 10^{-6}$ cm/sec)	30% to 50% clay (10^{-4} to 10^{-6} cm/sec)	15% to 30% clay (10^{-2} to 10^{-4} cm/sec)	0% to 15% clay ($<10^{-2}$ cm/sec)	8

Table XXIII (H-1) Continued

B-3 Potential for Ground-Water Contamination--Continued

Rating Factors	Rating Scale Levels 0	1	2	3	Multiplier
Subsurface flows	Bottom of site greater than 5 feet above high ground-water level	Bottom of site occasionally submerged	Bottom of site frequently submerged	Bottom of site located located below mean ground-water level	8
Direct access to ground water (through faults, fractures, faulty well casings, subsidence, fissures, etc.)	No evidence of risk	Low risk	Moderate risk	High risk	8

IV. WASTE MANAGEMENT PRACTICES CATEGORY

A. This category adjusts the total risk as determined from the receptors, pathways, and waste characteristics categories for waste management practices and engineering controls designed to reduce this risk. The total risk is determined by first averaging the receptors, pathways, and waste characteristics subscores.

B. Waste Management Practices Factor

The following multipliers are then applied to the total risk points (from A):

Waste Management Practice	Multiplier
No containment	1.0
Limited containment	0.95
Fully contained and in full compliance	0.10

Guidelines for fully contained:

Landfills:

- o Clay cap or other impermeable cover
- o Leachate collection system
- o Liners in good condition
- o Adequate monitoring wells

Surface Impoundments:

- o Liners in good condition
- o Sound dikes and adequate freeboard
- o Adequate monitoring wells

Spills:

- o Quick spill cleanup action taken
- o Contaminated soil removed
- o Soil and/or water samples confirm total cleanup of the spill

Fire Protection Training Areas:

- o Concrete surface and berms
- o Oil/water separator for pretreatment of runoff
- o Effluent from oil/water separator to treatment plant

General Note: If data are not available or known to be complete the factor ratings under items I-A through I, III-B-1, or III-6-3, then leave blank for calculation of factor score and maximum possible score.

AF

HAZARDOUS ASSESSMENT RATING FORM

NAME OF SITE ______________________________

LOCATION ______________________________

DATE OF OPERATION OR OCCURRENCE ______________________________

OWNER/OPERATOR ______________________________

COMMENTS/DESCRIPTION ______________________________

SITE RATED BY ______________________________

I. RECEPTORS

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
A. Population within 1,000 feet of site		4		
B. Distance to nearest well		10		
C. Land use/zoning within 1 mile radius		3		
D. Distance to reservation boundary		6		
E. Critical environments within 1 mile radius of site		10		
F. Water quality of nearest surface water body		6		
G. Ground water use of uppermost aquifer		9		
H. Population served by surface water supply within 3 miles downstream of site		6		
I. Population served by ground-water supply within 3 miles of site		6		
		Subtotals	____	____

Receptors subscore (100 X factor score subtotal/maximum score subtotal) ____

II. WASTE CHARACTERISTICS

A. Select the factor score based on the estimated quantity, the degree of hazard, and the confidence level of the information.

1. Waste quantity (S = small, M = medium, L = large) ____
2. Confidence level (C = confirmed, S = suspected) ____
3. Hazard rating (H = high, M = medium, L = low) ____

Factor Subscore A (from 20 to 100 based on factor score matrix) ____

B. Apply persistence factor
Factor Subscore A X Persistence Factor = Subscore B

____ X ____ = ____

C. Apply physical state multiplier

Subscore B X Physical State Multiplier = Waste Characteristics Subscore

____ X ____ = ____

Fig. 4. (H-2) Rating Form

III. PATHWAYS

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score

A. If there is evidence of migration of hazardous contaminants, assign maximum factor subscore of 100 points for direct evidence or 80 points for indirect evidence. If direct evidence exists then proceed to C. If no evidence or indirect evidence exists, proceed to B.

Subscore ______

B. Rate the migration potential for 3 potential pathways: surface water migration, flooding, and ground-water migration. Select the highest rating, and proceed to C.

1. Surface water migration

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
Distance to nearest surface water		8		
Net precipitation		6		
Surface erosion		8		
Surface permeability		6		
Rainfall intensity		8		

Subtotals ______ ______

Subscore (100 X factor score subtotal/maximum score subtotal) ______

2. Flooding

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
Flooding		1		

Subscore (100 x factor score/3) ______

3. Ground-water migration

Rating Factor	Factor Rating (0-3)	Multiplier	Factor Score	Maximum Possible Score
Depth to ground water		8		
Net precipitation		6		
Soil permeability		8		
Subsurface flows		8		
Direct access to ground water		8		

Subtotals ______ ______

Subscore (100 x factor score subtotal/maximum score subtotal) ______

C. Highest pathway subscore.

Enter the highest subscore value from A, B-1, B-2 or B-3 above.

Pathways Subscore ______

IV. WASTE MANAGEMENT PRACTICES

A. Average the three subscores for receptors, waste characteristics, and pathways.

Receptors ______
Waste Characteristics ______
Pathways ______

Total ______ divided by 3 = ______
Gross Total Score

B. Apply factor for waste containment from waste management practices

Gross Total Score X Waste Management Practices Factor = Final Score

______ X ______ = ______

Fig. 4. (H-2) Continued

APPENDIX B: Data Base of Site Characteristics

Key for Data Base

Column (s)	Variable
1-4	Site Identifier
5-6	General Disposal 10-Below Ground 01-Above Ground 00-Other or Both
7-10	Phase I Report Date
11	Population within 1,000 feet of site
12	Distance to nearest well
13	Land Use/Zoning within 1 Mile
14	Distance to Base Boundary
15	Critical Environments within 1 Mile
16	Surface Water Quality in Area
17	Ground Water Use of Upper Aquifer
18	Population Served by Surface Water Supply
19	Population Served by Ground Water Supply
20	Waste Quantity
21	Confidence Level
22	Hazard Rating
23-24	Persistence
25-27	Physical State
28-29	Evidence of Contamination 10-Direct Evidence 01-Indirect Evidence 00-No Evidence
30	Distance to Nearest Surface Water
31	Net Precipitation
32	Surface Erosion
33	Surface Permeability
34	Rainfall Intensity
35	Flooding
36	Depth to Ground Water
37	Net Precipitation
38	Soil Permeability
39	Subsurface Flows
40	Direct Access to Ground Water
41-43	Waste Management Practices
44	Phase II Site Classification 0-No Action 1-Long Term Monitoring 2 Required Action

45-48	Phase II Report Date
(49-59)	Type of Site
49	1-Landfill
50	1-Fire Training Area
51	1-Spill Site
52	1-Waste Disposal
53	1-Drum Disposal
54	1-Radioactive Waste Area
55	1-Leaking Underground Storage Facilities
56	1-Oil/Water Separator
57	1-Waste Treatment Plant
58	0-Munitions Disposal
59	1-Other
(60-67)	General Hazardous Waste Class
60	1-Volatile Organics
61	1-Petroleum, Oils and Lubricants
62	1-Heavy Metals
63	1-Radiological Substances
64	1-Polychlorinated Biphenyls
65	1-Sludge
66	0-Munitions
67	1-Other
68	Highest Pathway (GW/SW) 1-Ground Water (GW) 0-Surface Water (SW)
69-71	Receptors Category Score
72-74	Waste Characteristics Category Score
75-77	Pathway Category Score
78-80	Final HARM Score

Bases Included in Research

Duluth IAP, MN
Eglin AFB, FL
Griffiss AFB, NY
Hancock Field, NY
Hill AFB, UT
Kelly AFB, TX
Langley AFB, VA
Luke AFB, AZ
MacDill AFB, FL
McGuire AFB, NJ
Robins AFB, GA
Tinker AFB, OK
Tyndall AFB, FL
Wright-Patterson AFB, OH

Data used in HARM

1 040901820413220010321310100002111302121010028503010000000001000010044050062052

2 040810820413030010322310100001111302121009508503000100000000000010044080048054

3 040701820423220010331209100012111302121010028503001000000001000000046045080057

4 040401820413120110332309100003101302121010028503010000000011000000046090055064

5 040301820433230010312310100103103302121010028503001000000000000010051060100070

6 040510820423220010321310100011111302121010028503100000000001000010046050080059

7 040610820403220010331310075013111302122010028503100000000000000010042053080058

8 020101820611332130332310100003001101020310008411000000000101000001051100042064

9 020200820613330130332310100003001101020310008411000100000011000001062100042068

10 090101810610332200011310100003211113212 100082080000100000000000011035040067047

11 090200810610232200011310100003211113212 100082080010000000000001001033040067047

12 090310810610332200011310100003211113213 100082081000000000000000011035040076050

13 020301820623330130332310100003001101020310028411010000000001000001064100042069

14 030101820701233100122310100012212203211210028410010000000011000001042080080067

15 030210820731233100121310100003201203222309528410100000000011000001049050081057

16 020401820633230130322310100003001101020310028411010000000001000001065080042062

17 030310820721233100121310100002201203222309528410100000000011000001047050081056

18 030410820721233100111310100003202203213310028410100000000010000001047040081056

19 040100820423222010332310075103111302123010028503000100000000000110057075100077

20 040210820433222010322310075103111302123010028503100000000000000010059060100073

21 090410810600222200012310100003211113213 10008208000100000000000011028060076055

22 090610810610232200011308100003211113212 10008208001000100001000001033032067044

23 090710810600222200011310100003211113213 10018208100000000000000011028040076048

24 090810810600222200011310100003211113213 10018208100000000000000011028040076048

25 0909108106002322000113101000032111113213 100182081000000000000000011031040076049

26 0910018106103222000113081000032111113212 100082080010000000010000010320320670 44

27 0911108106103222000113081000032111113212 1000820800000010000100000103503206704 5

28 0905018106103322000113081000032111113212 1000820800100000000100000103503206704 5

29 1001108211132311103323091000032202032312100284101000000000000000011056090074073

30 1002108211232311103323080750032102132322100284101000000000000000011058060081066

31 1003108211333311203223080750032102032322100284101000000000000000011067048080065

32 1004108211030301100323091000022102022301100184100001000000001000001035090053059

33 1005018211332301103113101000132102032301100234100000000000100000011055040080058

34 1006008211232311103222091000032102032300100284100000101000010000100538054056056

35 1007018211112201103223091000022002032300100084100100000000010000110360720 53054

36 1008008211332201103223080750032002032301100284100001000000001000101052048060053

37 1009108211132311103113081000032002032311100284101000000000000000011056032067052

38 0501008204132200303122101000030012010200100085050000000000100010000054050042049

39 0502008204132200303112100500030012010200100085050000000000100010000054015042037

40 0503108204232300303223100750030312010200100285051000000000000000110059045065056

41 0504108204132200303113101000030012010200100185051000000000000000100540400420 45

42 0505108204232200303113101000030212310200100185051000000000000000100560401000 65

43 0506108204132200303123101000030312010200100185051000000000000000100540600650 60

44 0507108204132300303313100751030112110200100185051000000000010001100570531000 70

45 0508108204132300303113101000030212010200100185051000000000000000100570400570 51

46 0509108204132300303323101000030012010200100085050001000000011000000057100046068

47 0510108204132200303323101000010012010200100085050001000000011100001054100028061

48 0601108202323331333323101000132322022113095285021000000000000000100881000810 85

49 0602108202323331333323090750122322022113095285021000000000000000100880680800 75

50 0603108202120333333223091000032102322313095285021000000000000000100850721000 81

51 06041082021203333332221007501323023323333095285021000000000000000010085045100073

52 0605108202222333333312090500032102032333095285021000000000000000011091023095066

53 0606108202031333333123090750032202032323095285021000000000000000011090041088069

54 0607108202121333333223090750022222022113095085021000000000000000010087054067066

55 060801820212133133322308100003200223230310018502010000000001000000080064074073

56 070110811113130030332310100102001102023009528409100000000101000110561001 00081

57 07021081111313003033231010010200110202000952840900010000000100001110561001 00081

58 0703108111131300303323101000120111020210095284091000000000100000110561000 80075

59 080100820203232300332310100003001332022009528407000100000010100000061100100083

60 080210820203232300321310100003001332021009528407100000000101000000610501 00067

61 080301820201230300322310100103002302011009528407000100000010100000039080100069

62 080401820201220300332310100012001302010309528407000100000010100000036100080072

63 080500820203230300322310100002001302020010018407000000000110100000050080043058

64 101010821112231110321208100003220203231110028410100000000000000001105103206 7050

65 101100821112231110332110075003200203230110028410000100000000000010105103806 0050

66 110100820301022111112310100011200203233310028410000100000000000001103806009 5064

67 110201820301222311122210100011210203230310028410010000000011000011048060080063

68 110301820301220311122210100011210203230009528410010000000011000011037060080056

69 110400820313230311111210100011200203231310028410000010000000100000105403008 1055

70 110500820302230111112208100001200203231310028410001000000000100000103904008 1053

71 12010081120123220032231007500310021313 1 1002840800000000010100000104706006 9059

72 120200811203030200322310075003100213131 1001840800000000010100010104306006 9057

73 120310811203020200311310100000100213131 1002840810000000000000001104004006 9050

74 120410811203010200311310100000100213131 1002840800010000000000001103704006 9049

75 120500811203010100312310075001100212130 1002840800010000000100010103304505 1043

76 120601811203030200322310100003100213130 1002840801000000000100001104308006 0061

77 120701811202020100312310100003100213130 10028408010000000000000011031060060050

78 130100811120232210222310100013100313133 10028409000000000101000011047080087071

79 130210811100122210231310100003100313131 10028409100000000010001001043070069061

80 130310811100022210231310100003100313131 10028409100000000011000011036070069058

81 130401811100020210222310100002100313130 10028409010000000001000011025080060055

82 130510811100022210221310100003100313131 10028409100000000011000011036050069052

83 0101108106032331201223091000132302032312100182121000000000000000011063072080072

84 0102108106132321201213091000012312022210100182121000000000000000010060045061055

85 0103108106132321201122091000012212032210100082121000000000000000010060045054053

86 0104008106222321201213081000032012032300100282120000100000010000000057040057051

87 0105108106112132201212091000022012022200100182121000000000000000010051036046044

88 010600810611210120121310100002211202230010028212000100000000001000031050054045

89 0107108106212121201113091000032212022200100082121000000000000000010044036069050

90 010800810601210120122308100001201202320010008212000000000100000011029064056050

91 140110811003222200331310100002300303333309508409100000000011001111054070100071

92 140210811001012100331310100011300303333009518409100000000011001111033070080058

93 140310811003220100331310100002300303332009508409100000010001000011040070072058

94 1404008110000121003123081000123103033330095084090000000000100000011028048080049

95 1405008110030322003113101000033003033320095084090000000000100000011054040072052

96 140610811003211100321310100003310303331010018409100000000011001010042050069054

97 1407008110030101003123101000123003033300100284090000000000100000011033060080058

98 1408008110020201003223101000033003023300100084090000000000100000010031080061057

99 1409008110032011003113101000033203033320100084090001000000000000010039040076052

100 14100081100321010031131010001130030333301001840900000000000100100001037040080052

101 1411008110032101003113101000013003033330100084090000000000100000011037040079052

102 14121081103121210031131010000130130332220950840910000000000000000111043040079051

103 14131031103121210031131010000130130332201000840910000000000000111043040065049

104 14141081100121010031110405000130030333100950840910000000000000011026003065029

105 141500811003220100311310100003300303333011001840900000000110000011040040065048

106 141600811001220100311310100011300303333301000840900000000100000011029040080050

107 14170081100221010031131010000130030233201000840900000000100000011031040065045

108 141810811003222200331310100002300303333330952840910000000011001111054070100071

Bibliography

1. Air Force Engineering and Services Center/Air Force Occupational Environmental Health Laboratory. Memorandum on Rating System Development from Air Force Contractors: Engineering Science and CH2M Hill. HQ AFESC/DEV, Tyndall AFB FL: 31 January 1984.

2. Air Force Installation Restoration Management Committee, Col Klingensmith, HQ USAF/LEEV Chairman. Minutes of 27-28 March 1985 Meeting. Washington DC, 12 April 1985.

3. "Assessing, Managing Risk: EPA Chief Asks Scientists' Help," _Chemical and Engineering News_, _61_:4 (June 1983).

4. Caldwell, Steve, Project Officer, Hazard Ranking System. Telephone interview. United States EPA, Office of Emergency and Remedial Response, Washington DC, 14 November 1984.

5. _Comparison of U.S. Air Force Hazard Assessment Rating Methodology (HARM) with U.S. Environmental Protection Agency Hazard Ranking System (HRS) at Four Air Force Bases Evaluated under the Phase I Installation Restoration Program_. Report prepared for USAF HQ AFESC/DEV, Tyndall AFB FL. Engineering Science, Atlanta GA, April 1983.

6. "Comprehensive Environmental Response, Compensation, and Liability Act of 1980," _Environmental Reporter_. Washington DC: Bureau of National Affairs, Inc., _71_: 0701-0717 (20 April 1984).

7. Cornelius, Lt Col Kenneth. Chief Environmental Division. Personal interviews. HQ USAF, Bolling AFB Washington DC, 24-25 July 1985.

8. "Current Developments," _Environmental Reporter_. Washington DC: Bureau of National Affairs, Inc., _15_: 1-32 (4 May 1984).

9. "Current Developments," _Environmental Reporter_. Washington DC: Bureau of National Affairs, Inc., _15_: 753-784 (14 September 1984).

10. Emory, William C. _Business Research Methods_ (Revised Edition). Homewood: Richard D. Irwin, Inc., 1980.

11. "Environmental Protection Agency National Oil and Hazardous Substances Pollution Contingency Plan under Comprehensive Environmental Response, Compensation, and Liability Act of 1980," Environmental Reporter. Washington DC: Bureau of National Affairs, Inc., 101: 1001-1058 (16 November 1984).

12. General Accounting Office. Efforts to Clean Up DOD-Owned Inactive Hazardous Waste Disposal Sites. Report series NSIAD-85-41. Gaithersburg MD: GAO Document Handling and Information Services Facility, 12 April 1985.

13. General Accounting Office. Status of Air Force Efforts to Deal with Ground Water Contamination Problems at McClellan lAir Force Base. Report Series NSIAD-84-37. Gaithersburg MD: GAO Document Handling and Information Services Facility, 29 November 1983.

14. Hanson, David. "Industry Examines EPA's Risk Approach to Regulating Chemicals," Chemical and Engineering News, 61: 12-13 (December 1983).

15. Hernandez, Dr. John W. "Work Together to Protect Our Health and Our Environment," Journal of Environmental Health, 45: 63-65 (September/October 1982).

16. Hohenemser, C. et al. "The Nature of Technological Hazard," Science, 220: 378-381 (April 1983).

17. Huber, Peter. "Discarding the Double Standard in Risk Regulation," Technology Review, 87: 10-14 (January 1984).

18. Installation Restoration Program Records Search for Air Force Plant 6, Georgia. HQ AFESC Contract F08637-80-G0010-CH2M Hill, Gainsville FL, March 1984.

19. Kahan, James P. How Psychologists Talk About Risk. Rand Paper Series P-6403. Rand Corporation, Santa Monica CA, October 1979.

20. Klecka, William R. Discriminant Analysis Series: Quantitative Applications in the Social Sciences. Beverly Hills: Sage Publications, 1981.

21. Levine, Saul. "Probabilistic Risk Assessment: Identifying the Real Risks of Nuclear Power," Technology Review, 87: 40-44 (February/March 1984).

22. MacKenzie, James J. "Finessing the Risks of Nuclear Power," Technology Review, 87: 34-38 (February/March 1984).

23. McNichols, C. W. An Introduction to Applied Multivariate Data Analysis. Course notes. Air Force Institute of Technology (AU), Wright-Patterson AFB OH, 1980.

24. Miller, Stanton. "Risks: How to Get More Science in Assessments," Environmental Science and Technology, 17: 199A-200A (May 1983).

25. Morrison, Allen. "Risk Analysis: Is EPA Changing the Rules?" Civil Engineering-ASCE, 52: 60-63 (October 1982).

26. Nie, Norman H. et al. Statistical Package for the Social Sciences. New York: McGraw-Hill, 1975.

27. Ruckelshaus, William D. "Risk in a Free Society a Reservoir of Trust," Vital Speeches of the Day, 50: 345-348 (April 1984).

28. Salem, Steven L. et al. Issues and Problems in Inferring a Level of Acceptable Risk. Rand Report R-2561-DOE. Rand Corporation, Santa Monica CA, August 1980.

29. "Science, Risk, and Public Safety," Science, 221: 1026-1028 (September 1983).

30. Solomon, Kenneth A. et al. Dealing with Uncertainty Arising Out of Probabilistic Risk Assessment. Rand Report R-3045 ORNL. Rand Corporation, Santa Monica CA, September 1983.

31. Tabachnick, Barbara G. and Linda S. Fidell. Using Multivariate Statistics. New York: Harper and Row, 1983.

VITA

Mr. Myron C. Anderson was born on 18 September 1946 in Mercer, Pennsylvania. He graduated from Mercer High School in 1965 and attended Pennsylvania State University from which he received the degree of Bachelor of Technology in Water Resources Engineering Technology in June 1974. After graduation he was employed by the U.S. Army Corps of Engineers, Baltimore District, Maryland. While there he went through the civil engineer training program and worked in the Water Quality Control Branch until 1977. Next he worked for the U.S. Army at Ft. Indiantown Gap, Pennsylvania as the Chief of the Environment and Energy Office. Then in 1980 he went to work for the Air Force Engineering and Services Center, Tyndall AFB, Florida and remained there until entering the School of Systems and Logistics, Air Force Institute of Technology, in May of 1984.

Permanent Address: 927 Goose Bayou Road
Lynn Haven, Florida 32444

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A161066

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GEM/LSY/85S-1

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION: School of Systems and Logistics

6b. OFFICE SYMBOL (If applicable): AFIT/LS

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, OH 45433

7b. ADDRESS (City, State and ZIP Code)

8a. NAME OF FUNDING/SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State and ZIP Code)

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE (Include Security Classification): See box 19

12. PERSONAL AUTHOR(S): Mr. Myron C. Anderson, GM-13

13a. TYPE OF REPORT: MS Thesis

13b. TIME COVERED: FROM ______ TO ______

14. DATE OF REPORT (Yr., Mo., Day): 1985 September

15. PAGE COUNT: 132

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.
13	02	
15	05	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Hazardous Waste, Migration, Hazard Assessment, Toxic Hazards, Waste Management, Contaminants, Waste Disposal, Air Force Facilities,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: VALIDATION OF AIR FORCE HAZARD ASSESSMENT RATING METHODOLOGY

Thesis Chairman: Richard L. Murphy
Assistant Professor of Quantitative Management

Approved for public release: IAW AFR 190-1.
LYNN E. WOLAVER
Dean for Research and Professional Development
Air Force Institute of Technology (ATC)
Wright-Patterson AFB OH 45433

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: Richard L. Murphy

22b. TELEPHONE NUMBER (Include Area Code): 513-255-8410

22c. OFFICE SYMBOL: AFIT/LSY

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

The Air Force uses the Hazard Assessment Rating Methodology (HARM) for the initial screening of uncontrolled hazardous waste sites in the Installation Restoration Program (IRP). This initial screening evaluates the potential health and environmental hazards associated with the site and determines if the site warrants further investigation. There is a definite need to properly evaluate these sites because investigative costs are high and it is important not to eliminate sites that need further investigation.

This ~~research~~ evaluated results from actual Phase II investigations. Discriminant analysis was used to improve the HARM model's ability to properly evaluate sites.

The results indicate that current HARM procedures correctly predict which sites need further investigation and which do not only 68% of the time. Through the application of classification techniques developed in this study, the predictive capability was increased by over twenty percentage points so that approximately 90% of the sites were correctly classified. Before the refined model is applied more Phase II data are needed for final evaluation and testing of the new model.

END

FILMED

12-85

DTIC